

1719 CTGAGCGGTACACACAGCCGCTGTGTGCGCTGGGTGGCCACTTGGATCTTACTGAT 1778
1796 CTGAGCGGTACACACAGCCGCTGTGTGCGCTGGGTGGCCACTTGGATCTTACTGAT 1755
1779 CCAGCATGTATGATCTGATGACACTTGGATCTTGGATCTTGGATCTTGGATCTT 1888
1756 CCAGCATGTATGATCTGATGACACTTGGATCTTGGATCTTGGATCTTGGATCTT 1813
1839 GTAGCGCCGCTGACCAAAACATCAAGGTGTGGGTGGGTGGGTGGGTGGGTGGGT 1898
1714 GTAGCGCCGCTGACCAAAACATCAAGGTGTGGGTGGGTGGGTGGGTGGGTGGGT 1732
1899 CCAGCATGTATGATCTGATGACACTTGGATCTTGGATCTTGGATCTTGGATCTT 1942
7373 CCAGCATGTATGATCTGATGACACTTGGATCTTGGATCTTGGATCTTGGATCTT 7416

RESULT 13
S79463
LOCATION M-sema F-a factor in neural network development [mice, neonatal
ACCESSION S79463
VERSION S79463.1 GI:1110598
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3503)
Inagaki, S., Furuyama, T. and Iwahashi, Y.
Identification of a member of mouse semaphorin family
JOURNAL J. Biol. Chem. 269(13):2692-2697, 1994
MEDLINE 16538892
PUBMED 7636991
REMARK
Genbank staff at the National Library of Medicine created this
entry [NCBI g1bseq 170690] from the original journal article.
This sequence comes from Fig. 1A.
Location/Qualifiers
1..3503
/organism="Mus sp."
/mol_type="mRNA"
/db_xref="taxon:10095"
1..3503
/gene="M-sema F"
/note="a factor in neural network development"
100..2604
/gene="M-sema F"
/note="semaphorin homolog; this sequence comes from Fig.
1A"
/codon_start=1
/protein_id="MAB3184.1"
/db_xref="GI:1110598"
translation="MAPRPAVLLAAGLGLTGAEEMNIVPRKTVSGSELYTVRR
FSQGIQDFLTLTEHSGLVAGAREALAFVSELELQGISSEAPREKIECTOK
GSGNTECFNFRFLQPNNSHLVCGTAPQKTYIMLFTLDRFEFGKCP
YPAGHETLVDGELYSATLNNFGLTEVILRYGTHHSITETLAEVLNRPVGS
APVPSVSGTDDDKIYFSESEAVETICSEOVYARARVCKDMDGARLQKMT
TEKARVCSADMTVYVQKAVHLRGSAMHNTFEVFGARMDNDLSVCEQOL
EIOIQVEPTRETSPOAKARVYDPPSPRGSCINMHRDNGTSLDPMNTIN
ETKRPIMEDQVPRGLRPLVANKNTNFHVADRPGGLGATVYVLFCTGDMLK
AVSLKFWIMVELOVDFDPEVSLVLSQSKVLFAGSSQSLVOLSLACTYRCDV
CYLADPYCANVNTSRCAVATSGRGSFLVOHVALDTSKMGKNGIKRVSIPNI
TVVSGDIYLPCHLSNLAHMTGSDLPAPQPSFLYDGLQALVMAKSRISG
PYRCSYECGTSLAASLYAVVAGSVTLERAPLEGLVWLAVVAGACIVLL
LVSLKRLREBELERKAKASERTVYPIELPREPASPPRRPRDELTMDVGYYS
DSLKVTPGHARQEGGPPSPPGIPGQPLPSPTLRLHGGGRNSNANGTVRLQJGGE
DRGSGHPPELDELRLRRLQROPLPDSNPRESSV"

BASE COUNT 726 a. 917 c 1057 g 803 t
ORIGIN
Query Match 59.2%; Score 1351; DB 10; Length 3503;
Best Local Similarity 84.7%; Pred. No. 1,1e-268;

Matches 1539; Conservative 0; Mismatches 271; Indels 6; Gaps 2;
127 AGGACAGTGGCCGTGAAGTCAAGGCCGGGGCGTGGCCATGCCCCACACTGGGCTGTC 186
61 AGGCGTGTGAGTGAAGTTTAGACCTTGGGTGTCTGCATGAGCCCAACACTGGGCTGTC 120
187 TGGCTGTGGACCAAGGCTGTGGGGCTGGGGATTTGGGGCTGAGGTGTGGGAACTT 246
121 TGGCTGTGGACCAAGGCTGTGGGGCTGGGGATTTGGGGCTGAGGTGTGGGAACTT 180
247 GTGCGGGTGAAGACATGCTTCTGGGAGACTGGCCAGGTAGTACGGGGTTCCTCCAG 306
181 GTGCGGGTGAAGACATGCTTCTGGGAGACTGGCCAGGTAGTACGGGGTTCCTCCAG 240
307 ACCGGCATCCAGGACTTCTGACACTGACGCTGACGAGAGCCCACTGGGCTTCTGACGTG 366
241 ACAGGATCCAGGACTTCTGACACTGACGCTGACGAGAGCCCACTGGGCTTCTGACGTG 300
367 GGGCGCCGAGAGGCGCTGTTGCTTCCCTTCAGCATGAGAGCCCTGGAGCTGCAAGAGGATC 426
301 GGGCGCCGAGAGGCGCTGTTGCTTCCCTTCAGCATGAGAGCCCTGGAGCTGCAAGAGGATC 360
427 TCCGCGGAGGCGCCCTGCGGAGAGAGACTGAGGTATCCAGAAAGGAGAAACACAG 486
361 TCTTGGAGGCTCCAGCTGAGAGAGAAATTGAATTATCCAGAAAGGAGAAACACAG 420
487 ACCGAGTCTTCAACTTCACTCCGCTTCTGACAGCCCTCAATAGCTTCCACTGTACGTC 546
421 ACCGAGTCTTCAACTTCACTCCGCTTCTGACAGCCCTCAATAGCTTCCACTGTACGTC 480
547 TGTGGCACTACGCTTCCAGCCCAAGTGCACCTACCTGCAAGTCCCTACCTTCACTT 606
481 TGTGGCACTACGCTTCCAGCCCAAGTGCACCTACCTGCAAGTCCCTACCTTCACTT 540
607 GAGCATGAGAGTGAAGATGAGAGAGGCAAGTGTCCCTATGACCCAGCTAAGGGCAT 666
541 GAGCATGAGAGTGAAGATGAGAGAGGCAAGTGTCCCTATGACCCAGCTAAGGGCAT 600
667 GCTGCTTCTTGTGATGTGAGCTGTACCTGCGCACACTCAACAACTTCTTGGGACG 726
601 ACCGACTCTTGTGAGAGTGTGAGCTGTACCTGCGCACACTCAACAACTTCTTGGGACG 660
727 GAACCATTTCTGCGGTGATGAGTGGGCGCCAGCACTGCAAGAGAGAGTACTGCGC 786
661 GAGCGGTATCTTCTGATGATGAGTGGGCGCCAGCACTGCAAGAGAGTACTGCGC 720
787 TTTGGCTCAAGAACTCACTTGTAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 846
721 TTTGGCTCAAGAACTCACTTGTAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
847 TTTGGCTCAAGAACTCACTTGTAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 906
781 TTTGGCTCAAGAACTCACTTGTAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
907 TTTGGCTCAAGAACTCACTTGTAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
841 TTTGGCTCAAGAACTCACTTGTAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
967 GAGCGGAGCTGCAAGAGAGTGAAGTGTGCTGCAAGAGGCGGCTGCAAGAGTGTGCTGCT 1026
901 GAGCGGAGCTGCAAGAGAGTGAAGTGTGCTGCAAGAGGCGGCTGCAAGAGTGTGCTGCT 960
1027 CGGAAGCTGCAAGAGTGTGCTGCAAGAGGCGGCTGCAAGAGTGTGCTGCTGCTGCTGCT 1086
961 CGGAAGCTGCAAGAGTGTGCTGCAAGAGGCGGCTGCAAGAGTGTGCTGCTGCTGCTGCT 1020
1087 TTTGGCTCAAGAACTCACTTGTAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1146
1021 TTTGGCTCAAGAACTCACTTGTAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
1147 GCGATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1206
1081 GCGATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140

THIS PAGE BLANK (USPTO)

Wed Aug 20 09:19:27 2003

us-10-002-05

1207 TACCATGAGAGAGCCAGAGAGTGGAGCCGCTACACTGACCCTGTACCCAGCCCTGGGCT 1266
1141 TACAGTANAGAGCCAGAGAGTGGGCGGCTATATAGTACCCGGTACCCAGCCCTGGGCT 1200
1267 GGCTCTGTCATTAACAACCTGGCTCGGCGGACGGCTACACAGCTCCCTGGAGCTACCC 1326
1201 GGTTCGTGTATCAACAACCTGGCACAGACAAATGGCTACACAGCTCCCTGGAGCTCGG 1260
1327 GACAACTCCCAACTGCTCAAGAGAGCCGCTGATGAGAGACAGTGGGCTCGG 1386
1261 GACAAACACCCCTCATCAAGAGAGACCCCTGATGAGAGACAGTGAAGCTCGG 1320
1387 TGAAGCCGCCCCCTGCTGCTGAGAGAGAGGACCAACTTACCCACTGCTGGCGGACCGG 1446
1321 TTGGGCCCCCTCTCTGTGAGAGAGAACTACTTACACACAGTGTGGCGGACAGG 1380
1447 GTTACAGAGCTGATGAGAGCCACTATACAGTCTGTCATTTGGCAGAGAGAGGCTGG 1506
1381 GTCCAGGGCTGTATGCTGCGCCACTATACAGTGTGTCATTTGTCAGAGAGATGCTGG 1440
1507 CTGCTCAAGGCTGTAGAGCTGGGCGCTGGGTTCACTGATTTGAGAGAGCTGAGCTTT 1566
1441 CTGCTGAAGGCTGTGAGGCTGGGCGCTGGATCCACATGTGTGAGAGAGTGCAGGTTT 1500
1567 GACCAAGAGCCCATGAGAGAGCTGTGCTATCTCAGAGCAGAGAGCTGCTTTGGCGGC 1626
1501 GACCAAGAGCCCATGAGAGAGTGTGCTGTCTCAGAGCAGAGAGAGTGTCTTTGCTGGC 1560
1627 TCCGCTCTCAGCTGTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1686
1561 TCCGCTCTCAGCTGTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
1687 GACTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1746
1621 GACTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
1747 G---CCGTTGGTGGCCACTTTGATCTTACTGATCCAGATGTGATGATGATGATGATGATG 1803
1681 GCCACACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
1804 TCAGGCAATTGCAACCTCCGCTGGAGTAAAGATACAGTCAAGGCCACTGCCAATAACATC 1863
1741 TCAAGATGTGTAAACAGATGCAATTA---AAAGTCAATCTATTTCCAAAGACATC 1797
1864 ACGGTGTGGGGGACAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1923
1798 ACCGTGTGTACAGCAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1857
1924 CCCGACTGCAACCCCG 1939
1858 GCCCACTGACCTTCG 1873

THIS PAGE BLANK (USPTO)

JOURNAL Patent: WO 0078802-A 13 28-DEC-2000;

Curagen Corporation (US)

Location/Qualifiers

FEATURES

source

1.. 2284

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

166..1956

/note="unnamed protein product"

/codon_start=1

/protein_id="CAC24872.1"

/db_xref="GI:12405795"

/translation="MAPRAYVLLAARLMLGICGLVWVNLVPRKTVSSGELATVYRR

FSQIGIDFLVLTLEPTGLVYGARELFAFSMEALEOGAISEAPVEKTEKIOK

GNNGTECFNFRLOPYNASHLYVCTYVNLFTLEHGFEDKGCPC

YDPAGHAGLIVDELVSATLNNLSTEPILIRNMGRPHSMKTEYLAEVNEPVG

AYPESVSGFDGDKVYFFPERAVESPCYEOVVAVARAKEDMGARLORKMT

TEIKARLASAPNOLYFNOLQAMHTLQDTSMHTNTPFCVROAKGMDVYLSICXYOL

BEIOKVTPKYEHEBAQKNDRTDPIPSRPGSCINNMHRHGTSSLELPDNLN

FKKIPDLMEVOYGRPMRSLVLYKGTNTHLVADVTGLDQATVYVLEITGDSMLK

AVSLGPMVHLIEELQLEQDEPMRSLVLSQSKLFPAGSRSLVOLPVADCIKRYSCAD

CVLARDPYCAMSVTNSRCVAVGHSGLIIOHVMTSIDISGLCNLGRSKIOSGLPRKNI

TVVAGTDVLPCHLSSNLALPDSNPESSV"

BASE COUNT 443 a 683 c 689 g 465 t 4 others

ORIGIN

Query Match 99.8%; Score 2280; DB 6; Length 2284;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCCCTCTACACTCCGCTGCTGATGTGGAACGGGGTTGGGGTTCTGCAGGCT 60
DB 1 GGGCCCTCTACACTCCGCTGCTGATGTGGAACGGGGTTGGGGTTCTGCAGGCT 60
QY 61 ATTGTCTGCGTGGGGAAGGGGACAGGCCGGGACCTCCGCTCGACGCCGCCG 120
DB 61 ATTGTCTGCGTGGGGAAGGGGACAGGCCGGGACCTCCGCTCGACGCCGCCG 120
QY 121 ACCAGACAGACAGTGGCTGAAGCTCAGAGCCGGGGCGTGGCCATGGCCCCACACTGG 180
DB 121 ACCAGACAGACAGTGGCTGAAGCTCAGAGCCGGGGCGTGGCCATGGCCCCACACTGG 180
QY 181 GCTGTCTGGCTGTGGACAGAGGCTGTGGGGCTGGGGATGGGGCTGAGTGTGGTGG 240
DB 181 GCTGTCTGGCTGTGGACAGAGGCTGTGGGGCTGGGGATGGGGCTGAGTGTGGTGG 240
QY 241 AACCTGTGCGCGTGAAGACAGTGTCTGTGGGAGCTGGCCACGGTAGTACGGCGTTC 300
DB 241 AACCTGTGCGCGTGAAGACAGTGTCTGTGGGAGCTGGCCACGGTAGTACGGCGTTC 300
QY 301 TCCGACAGCCGGCATCTGACAGACTGTGACACTGACGCTGACGGACCCCACTGGGCTTCG 360
DB 301 TCCGACAGCCGGCATCTGACAGACTGTGACACTGACGCTGACGGACCCCACTGGGCTTCG 360
QY 361 TACGTGGGGGCCGAGAGAGCCCTGTGTGCTTCAAGATGAGAGCCCTGAGAGTGAAGA 420
DB 361 TACGTGGGGGCCGAGAGAGCCCTGTGTGCTTCAAGATGAGAGCCCTGAGAGTGAAGA 420
QY 421 GCGATCTCTGTGGAGGCCCGCTGGAGAGAGACTGAGTGTATTCAGAAAGGAGAAAC 480
DB 421 GCGATCTCTGTGGAGGCCCGCTGGAGAGAGACTGAGTGTATTCAGAAAGGAGAAAC 480
QY 481 AACGACAGCCGAGTCTTCAACTTCATCCGCTTCTGACAGCCCTACAAATCTCCCACTCG 540
DB 481 AACGACAGCCGAGTCTTCAACTTCATCCGCTTCTGACAGCCCTACAAATCTCCCACTCG 540
QY 541 TACGCTGTGGACCTAGAGCCCTTCAAGCCCAAGTCACTAGTCAAACTGTCTACCTTC 600
DB 541 TACGCTGTGGACCTAGAGCCCTTCAAGCCCAAGTCACTAGTCAAACTGTCTACCTTC 600
QY 601 ACTTGGAGCATGAGAGATTGAAGTGGAGGAGGCAAGTGTCCATGAGCCAGCTAAG 660
DB 601 ACTTGGAGCATGAGAGATTGAAGTGGAGGAGGCAAGTGTCCATGAGCCAGCTAAG 660

QY 661 GGCATGCTGGCCCTTCTGTGTGATGAGTGTACTGTGGCCACACTCAAACTTCCTG 720
DB 661 GGCATGCTGGCCCTTCTGTGTGATGAGTGTACTGTGGCCACACTCAAACTTCCTG 720
QY 721 GGCAGGAAACCATTTATCTCGGTAAATGAGGGGCCACCACTCCATGAAGACAGTAC 780
DB 721 GGCAGGAAACCATTTATCTCGGTAAATGAGGGGCCACCACTCCATGAAGACAGTAC 780
QY 781 CTGGCCTTTGGCTCAAGACACTCTTGTAGGCTGCGCATGATGATGAGTGTG 840
DB 781 CTGGCCTTTGGCTCAAGACACTCTTGTAGGCTGCGCATGATGATGAGTGTG 840
QY 841 GGCAGCTTACAGGGGGACAGACAGACAGTCTACTTCTTCTTCAAGAGCGGGCATGGAG 900
DB 841 GGCAGCTTACAGGGGGACAGACAGACAGTCTACTTCTTCTTCAAGAGCGGGCATGGAG 900
QY 901 TCCGACTGTATAGCCGAGCAGAGTGGTGGCTCGTGTGGCCCGTGTCTGCAAGGGCATATG 960
DB 901 TCCGACTGTATAGCCGAGCAGAGTGGTGGCTCGTGTGGCCCGTGTCTGCAAGGGCATATG 960
QY 961 GGGGGGCGACAGACCTGCGAGAGGAGTGGACAGCTTCCTGAAGCGGGGGTGGCATGC 1020
DB 961 GGGGGGCGACAGACCTGCGAGAGGAGTGGACAGCTTCCTGAAGCGGGGGTGGCATGC 1020
QY 1021 TCTGCCCGCACTGGCAGCTTACTTCAACAGCTGACAGGCGATGCAACCTGACAGAC 1080
DB 1021 TCTGCCCGCACTGGCAGCTTACTTCAACAGCTGACAGGCGATGCAACCTGACAGAC 1080
QY 1081 ACCTCTGGCACAACACACACTCTTGTGGGGTTTTCAGACACAGTGGGGTACATGTAC 1140
DB 1081 ACCTCTGGCACAACACACACTCTTGTGGGGTTTTCAGACACAGTGGGGTACATGTAC 1140
QY 1141 CTGTGCGCATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
DB 1141 CTGTGCGCATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
QY 1201 AAGGATACCATGAGAGAGCCAGAGAGTGGAGCCGCTACACTGACCTGTACCCAGCCT 1260
DB 1201 AAGGATACCATGAGAGAGCCAGAGAGTGGAGCCGCTACACTGACCTGTACCCAGCCT 1260
QY 1261 GGGCGGTGCTGTGATTAACAATGCGGCGGACAGGCTTACACACAGCTCCCTGGAG 1320
DB 1261 GGGCGGTGCTGTGATTAACAATGCGGCGGACAGGCTTACACACAGCTCCCTGGAG 1320
QY 1321 CTACCGACACATCTCTCAACTTCTGCAAGAAACACCCCTGATGAGAGACAGTGGAG 1380
DB 1321 CTACCGACACATCTCTCAACTTCTGCAAGAAACACCCCTGATGAGAGACAGTGGAG 1380
QY 1381 CCTCGGTGAGCCGCGCCCTGCTGTGAAGAAAGGACACCACTTACCCACTGTGGCC 1440
DB 1381 CCTCGGTGAGCCGCGCCCTGCTGTGAAGAAAGGACACCACTTACCCACTGTGGCC 1440
QY 1441 GACCGGTTACAGAGACTTATGAGGACACTTATGAGGCTGTCTGTTGACAGAGAGAC 1500
DB 1441 GACCGGTTACAGAGACTTATGAGGACACTTATGAGGCTGTCTGTTGACAGAGAGAC 1500
QY 1501 GCGTGGCTCTCAAGGCTGTGAGCCCTGGGGCTGAGTGTACCTGATTTAGAGAGCTGCAG 1560
DB 1501 GCGTGGCTCTCAAGGCTGTGAGCCCTGGGGCTGAGTGTACCTGATTTAGAGAGCTGCAG 1560
QY 1561 CTGTTTGAACAGAGCCCATGAGAAAGCTGTGTATCTCAGAGCAAGAGTGTCTTT 1620
DB 1561 CTGTTTGAACAGAGCCCATGAGAAAGCTGTGTATCTCAGAGCAAGAGTGTCTTT 1620
QY 1621 GCGGCTCCGCTCTCAGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1680
DB 1621 GCGGCTCCGCTCTCAGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1680
QY 1681 TGTGACAGTGTGTCTGCGCGGGAGCCCTATTGCGCTGGAGGCTCAACACAGCCGC 1740
DB 1681 TGTGACAGTGTGTCTGCGCGGGAGCCCTATTGCGCTGGAGGCTCAACACAGCCGC 1740

[illegible]

LOCUS	AX060303	2155 bp	DNA	linear	PAT 22-JAN-2001
DEFINITION	Sequence 11 from Patent WO00/78802.				
ACCESSION	AX060303				
VERSION	AX060303.1	GI:12405792			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 Shimkets,R.A., Fernandes,E., Vernet,C., Yang,M., Boldog,F.L. and Herrmann,J.L.				
TITLE	Secreted polypeptides and corresponding polynucleotides				
JOURNAL	Patent: WO 00/78802-A 11 28-DEC-2000;				
FEATURES	Curagen Corporation (US)				
source	Location/Qualifiers				
	1..2155				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	166..1938				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAC24871.1"				
	/db_xref="GI:12405793"				
	/translation="MAPHWAVLAAARLIMGICGAEVWMTVPKTVSSGLIAVWR				
	ESQICIDFLITLTPRGLLYGARGRLPFSMALELOCAISMSEAPVEKTKICIQE				
	GNNOTECFNTIRLQPTNASHLYVCGTFAQPCVYNNMLFTLEGEEDGKCE				
	YDPAGHAGGLLVDELVSATINNELGEPDILIRNMGPHSMKTEYLAFMLEPFRVSS				
	AYVPSVSGFTGDDDKLYVFFERRERVESDCAAEQVAVAVACGMDGARTLQKRYT				

.BASE COUNT		444	a	614	c	645	g	452	t
ORIGIN									
Query Match	Best Local Similarity	92.6%;	Score 1887.4;	DB 6;	Length 2155;				
Matches 1914;	Conservative	0;	Mismatches	6;	Indels	3;	Gaps	2;	
OY	1	CGGCCCTTCTC	ACTCTCTGCCCTGCTGATGTGGAACGGGGTTGGGGTTCTGCAGGGCT	60					
Db	1	CGGCCCTTCTC	ACTCTCTGCCCTGCTGATGTGGAACGGGGTTGGGGTTCTGCAGGGCT	60					
OY	61	ATTGTCGTGGCTGGGGGAAAGGGACAGCCGGGACCGGGACCTCCGCTGCAACCGGGCCG	120						
Db	61	ATTGTCGTGGCTGGGGGAAAGGGACAGCCGGGACCGGGACCTCCGCTGCAACCGGGCCG	120						
OY	121	ACCAGAGGACAGCTGGGCTGAAAGCTAGAGGCGGGGGGCTGGCGCATGGCCCCACACTGG	180						
Db	121	ACCAGAGGACAGCTGGGCTGAAAGCTAGAGGCGGGGGGCTGGCGCATGGCCCCACACTGG	180						
OY	181	GCTGTCTGTGCTCTGTCGACAGAGCTGTGGGGCTTGGGCAATGGGGCTGAGGTGTGTGG	240						
Db	181	GCTGTCTGTGCTCTGTCGACAGAGCTGTGGGGCTTGGGCAATGGGGCTGAGGTGTGTGG	240						
OY	241	AACCTTGTCCCGCGTGAAGACAGTGTCTTCTGGGGAGCTGGCCACGGTATGACAGCGGGTTC	300						
Db	241	AACCTTGTCCCGCGCGTGAAGACAGTGTCTTCTGGGGAGCTGGCCACGGTATGACAGCGGGTTC	300						
OY	301	TCCCAAGACCGGCAATCCAGACTTCCCTGACACTGAGAGCTGAGGAGGCCACCTGGGCTTCTG	360						
Db	301	TCCCAAGACCGGCAATCCAGACTTCCCTGACACTGAGAGCTGAGGAGGCCACCTGGGCTTCTG	360						
OY	361	TACGTGGGCGCCCGAGAGAGCCCTGTTTGGCTTCACACATGAGAGCCCTGGAGCTGCAAGGA	420						
Db	361	TACGTGGGCGCCCGAGAGAGCCCTGTTTGGCTTCACACATGAGAGCCCTGGAGCTGCAAGGA	420						
OY	421	GCGATCTCCTGGGAGGCCCCCGTGGAGAAAGACTGATGTATCCAGAAAGGAGAAAC	480						
Db	421	GCGATCTCCTGGGAGGCCCCCGTGGAGAAAGACTGATGTATCCAGAAAGGAGAAAC	480						
OY	481	AACCAAGACCGAATGCTTCAACTTATCCGCTTCTGCAAGCCCTACAAATGGCTCCCACTGG	540						
Db	481	AACCAAGACCGAATGCTTCAACTTATCCGCTTCTGCAAGCCCTACAAATGGCTCCCACTGG	540						
OY	541	TACGCTGTGGGACCTACAGCTTCCAGCCCAAGTCACTACGTCAACATGCTCACTTC	600						
Db	541	TACGCTGTGGGACCTACAGCTTCCAGCCCAAGTCACTACGTCAACATGCTCACTTC	600						
OY	601	ACTTTGGAGCATGAGAGTTTGAAGATGGGAAAGGCAAGTGTCCCTATGACCCAGCTAAG	660						
Db	601	ACTTTGGAGCATGAGAGTTTGAAGATGGGAAAGGCAAGTGTCCCTATGACCCAGCTAAG	660						
OY	661	GCGCATGCTGGGCTTCTTGTGATGTGAGTGTACTGTGGCCACACTCAACAATTTCTG	720						
Db	661	GCGCATGCTGGGCTTCTTGTGATGTGAGTGTACTGTGGCCACACTCAACAATTTCTG	720						
OY	721	GGCACGAAACCCATTATCTCGCTTAACATGAGGGGCCCCACCACTCCATGAAGCAAGATAC	780						
Db	721	GGCACGAAACCCATTATCTCGCTTAACATGAGGGGCCCCACCACTCCATGAAGCAAGATAC	780						
OY	781	CTGGCCCTTTTGGCTCAAGAACCTCACTTGTAGGCTGTGCTATGTACTAGAGTGTG	840						
Db	781	CTGGCCCTTTTGGCTCAAGAACCTCACTTGTAGGCTGTGCTATGTACTAGAGTGTG	840						
OY	841	GCGAGCTTACGAGGGGACGACGACAAGGTCTACTTCTTACAGGAGCGGGGACGTGGAG	900						
Db	841	GCGAGCTTACGAGGGGACGACGACAAGGTCTACTTCTTACAGGAGCGGGGACGTGGAG	900						

TEIKRIACSAANMOLYTRNOLOAMHTLODTSNHTTFFGCVROAKGDMYLAISCEYQ
BELIQVREBPKEVETHEAKMDRYTDVPSPRPSCINNMHRRHETYSLEIPNLT
EYKRPPLMEEOYGPMPRSRLPLVKKTNTHTLVADGTIDATATYGLDGMLEK
ANSLPWWHLLELOLFDQEPMRSLVLSOSKLLFAGRSRLOVLPVADCKIRYSAL
CYLADPYCAEMSVNTRCAVAGHGFSLLIHVMN SDPSGICNINLGSKVRPTPRNII
VYAGIDVLPHLLSSTMPRGSVYF*

```

Oy 901 TCCACTCTATGCGGACAGAGTGGTGGCTGCTGTGGCCCTGTGTGCAAGGCGCATATG 960
Db 901 TCCGACTCTATGCGGACAGAGTGGTGGCTGCTGTGGCCCTGTGTGCAAGGCGCATATG 960
Oy 961 GGGGCGGACGAGACCTCTCAGAGAGAGTGGACACAGTCTCTGAAAGCGCGGCTGGCATGC 1020
Db 961 GGGGCGGACGAGACCTCTCAGAGAGAGTGGACACAGTCTCTGAAAGCGCGGCTGGCATGC 1020
Oy 1021 TCTGCCCCGAACTGGCAGCTCTACTTCAACGAGCTGACGCGGATGACACACCTGGCAGAC 1080
Db 1021 TCTGCCCCGAACTGGCAGCTCTACTTCAACGAGCTGACGCGGATGACACACCTGGCAGAC 1080
Oy 1081 ACCCTCTGAGCAACACACCTCTCTTGGGGTTTTCAAGCAGCAGTGGGGTGACATGTAC 1140
Db 1081 ACCCTCTGAGCAACACACCTCTCTTGGGGTTTTCAAGCAGCAGTGGGGTGACATGTAC 1140
Oy 1141 CTGTGCGGCACTGTGATACCACTGTGAGAGAGATCCAGCGGGTGTGGAGCGCCCTAT 1200
Db 1141 CTGTGCGGCACTGTGATACCACTGTGAGAGAGATCCAGCGGGTGTGGAGCGCCCTAT 1200
Oy 1201 AAGGAGTACATGAGGAAAGCCGAGAGTGGAGCCGCTACACTGACCTGTACCCAGCCCT 1260
Db 1201 AAGGAGTACATGAGGAAAGCCGAGAGTGGAGCCGCTACACTGACCTGTACCCAGCCCT 1260
Oy 1261 GCGGCTGGCTGTGCTATTAACACATGCGATGGGGCCACGCGCTACACCACTCCCTGGAG 1320
Db 1261 GCGGCTGGCTGTGCTATTAACACATGCGATGGGGCCACGCGCTACACCACTCCCTGGAG 1320
Oy 1321 CTACCCGCAACATCTCTCAACTGCTCAAGAAAGCAGCGCTGATGAGAGAGAGTGGGG 1380
Db 1321 CTACCCGCAACATCTCTCAACTGCTCAAGAAAGCAGCGCTGATGAGAGAGAGTGGGG 1380
Oy 1381 CCGTGGTGGAGCCGCGCCCTCTCTCTGTAAGAAAGGAGCCAACTTACCCACTGGTGGCC 1440
Db 1381 CCGTGGTGGAGCCGCGCCCTCTCTCTGTAAGAAAGGAGCCAACTTACCCACTGGTGGCC 1440
Oy 1441 GACCGGTTTACAGAGCTGATGAGAGCCAGCTATACAGTGTGCTCATTTGGGACAGGAAAC 1500
Db 1441 GACCGGTTTACAGAGCTGATGAGAGCCAGCTATACAGTGTGCTCATTTGGGACAGGAAAC 1500
Oy 1501 GCGTGGCTGTCAAGAGCTGTGAGAGCTGGGGCCCTGGGCTTCACTGATTTAGAGAGCTGAG 1560
Db 1501 GCGTGGCTGTCAAGAGCTGTGAGAGCTGGGGCCCTGGGCTTCACTGATTTAGAGAGCTGAG 1560
Oy 1561 CTGTTTACAGAGAGCCCATGAGAAAGCTGTGCTATCTCAGAGCAAGAAAGCTGCTCTTT 1620
Db 1561 CTGTTTACAGAGAGCCCATGAGAAAGCTGTGCTATCTCAGAGCAAGAAAGCTGCTCTTT 1620
Oy 1621 GCGGCTGGCTGTCAAGAGCTGTGAGAGCTGGGGCCCTGGGCTTCACTGATTTAGAGAGCTGAG 1680
Db 1621 GCGGCTGGCTGTCAAGAGCTGTGAGAGCTGGGGCCCTGGGCTTCACTGATTTAGAGAGCTGAG 1680
Oy 1681 TGTGACAGATGTGCTCTGCGCGGAGCCCTATTGCGGCTGAGAGGCTCAACACAGCCGC 1740
Db 1681 TGTGACAGATGTGCTCTGCGCGGAGCCCTATTGCGGCTGAGAGGCTCAACACAGCCGC 1740
Oy 1741 TGTGTGGCGGTGGTGGGCACTTTGATCTTTACTGATCCAGCATGTGATGACCTCGGAC 1800
Db 1741 TGTGTGGCGGTGGTGGGCACTTTGATCTTTACTGATCCAGCATGTGATGACCTCGGAC 1800
Oy 1801 ACTTCAGGCACTTTGCAACCTCGGTGGGCACTGAAAGTACAGTACAGCCCACTCCAAAAAC 1860
Db 1801 ACTTCAGGCACTTTGCAACCTCGGTGGGCACTGAAAGTACAGTACAGCCCACTCCAAAAAC 1860
Oy 1861 ATCAGAGTGTGCGGCGACAGAGCTGTGCTGCGGCTGCACTCTCTCCAACTTGGCC 1920
Db 1861 ATCAGAGTGTGCGGCGACAGAGCTGTGCTGCGGCTGCACTCTCTCCAACTTGGCC 1920
Oy 1921 CTG 1923
Db 1918 CGG 1920

```

```

RESULT 3
AX060313
LOCUS 2156 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 21 from Patent WO078802.
ACCESSION AX060313
VERSION AX060313.1 GI:12405802
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Shinkets,R.A., Fernandes,E., Vernet,C., Yang,M., Boldog,F.L. and
Herrmann,J.L.
TITLE Secreted polypeptides and corresponding polynucleotides
JOURNAL Patent: WO 0078802-A 21 28-DEC-2000;
Curagen Corporation (US)
FEATURES
source location/Qualifiers
1..2156
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
156..2040
/codon_start=1
/codon="unnamed protein product"
/db_xref="GI:12405803"
/translation="MAPMAVWILAAKRLMGIGIAEYVMMNLVPRKTVSSGELATVVR
FSQNGIDELTLEPTGLIYVGRALFAPSMALELQASIAEPAVEKEDKRC
KNNOTREFNTRPLQPYNNASHLYCGTYAOPKTYNNMLTLEHEFEDKRC
YDPAKHAGLIVDELVSATLNNFATPEITLIRNRPKHSKTEYLAWLPEHVS
ATPESVGSFTGDDKVFYFFRERAVESDCAEDVAVAVARCKDGMGANTLDRKT
TEFLARLAKSA.PNNOLYFNOLAMHTLDDTSMHNTFEGVQADGMDYLAICEYOL
EIOIRVEGAPYKEHEBAOKMDRYDPPSPRPGSCIINMHRHGTSLSELPOWILN
FVSKHPLMEEOVGPRMSRPLLVKGTNTFHVADRVGLDQATVTVLFITGDEWLK
AVKGRPVHVLIELEOLFDEPFRMSLVLSOSKLLFAGRSLOLYLPVADCIKYRSCAD
CVIADRPYCAHSVTSRCAVANGHSGLILIOHVTSDPSGICNLNGSKRYPTPRNTT
VYAGIDLVLPCHLSITMPRGSVFTCLLPVQWERLECKRKRKMGACGQWHTLFP
"
BASE COUNT 444 a 617 c 645 g 450 t
ORIGIN
Query Match 82.3%; Score 1879.4; DB 6; Length 2156;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1909; Conservative 0; Mismatches 11; Indels 3; Gaps 2;
Oy 1 CGGCCCTCTCAGACTCTGCGCTGCTGATGAGAAAGGGGTTGGGTTGCGAGGCT 60
Db 1 CGGCCCTCTCAGACTCTGCGCTGCTGATGAGAAAGGGGTTGGGTTGCGAGGCT 60
Oy 61 ATTGTCTGCGCTGGGGAAGGAGGAGCCGCGGACCTCCGCTGCGACGCCGCGC 120
Db 61 ATTGTCTGCGCTGGGGAAGGAGGAGCCGCGGACCTCCGCTGCGACGCCGCGC 120
Oy 121 ACCAGCAGACAGCTGGGCTGAAAGCTCAGAGCCGGGGCTGGCCATGGCCACACCTGG 180
Db 121 ACCAGCAGACAGCTGGGCTGAAAGCTCAGAGCCGGGGCTGGCCATGGCCACACCTGG 180
Oy 181 GCTGTGAGCTGTGGGCAAGAGCTGTGGGGCCCTGGGCACTTGGGGCTGAGGTGTGGG 240
Db 181 GCTGTGAGCTGTGGGCAAGAGCTGTGGGGCCCTGGGCACTTGGGGCTGAGGTGTGGG 240
Oy 241 AACCTTGTGCGCGGTAGACAGTGTCTTCTGGGAGCTGGCCACGATGATACGGCGCTTC 300
Db 241 AACCTTGTGCGCGGTAGACAGTGTCTTCTGGGAGCTGGCCACGATGATACGGCGCTTC 300
Oy 301 TCCGAGACCGGCGATTCAGAGACTTCTGTACACTGACGCTGAGAGGAGCCCACTGGCTTTC 360
Db 301 TCCGAGACCGGCGATTCAGAGACTTCTGTACACTGACGCTGAGAGGAGCCCACTGGCTTTC 360
Oy 361 TACGTGGGCGCGGAGAGAGCCCTGTTTGCCTTCAGCATGAGAGCCCTGGAGCTGCAAGA 420
Db 361 TACGTGGGCGCGGAGAGAGCCCTGTTTGCCTTCAGCATGAGAGCCCTGGAGCTGCAAGA 420

```

OY	421	GC	GCATCTCTCTGGGAGGCCCCCGCTGGAGAAAGACTGATGTATTCAGAAAGGAGAAC	480
Db	421	GC	GATCTCTCTGGGAGGCCCCCGCTGGAGAAAGACTGATGTATTCAGAAAGGAGAAC	480
OY	481	AA	CCAGACCCGAGTGCCTTCAACTTATCCGCTCTCGAGGCCCTCAATGCCCTCCACTG	540
Db	481	AA	CCAGACCCGAGTGCCTTCAACTTATCCGCTCTCGAGGCCCTCAATGCCCTCCACTG	540
OY	541	TAC	GTCTGTGGACCTACGCCCTTCAGGCCCAAGTGACCTACGTCACATGCTCACCTTC	600
Db	541	TAC	GTCTGTGGACCTACGCCCTTCAGGCCCAAGTGACCTACGTCACATGCTCACCTTC	600
OY	601	AC	TTTGAGGAGATGAGAGTTGGAAGATGGGAAGGGCAAGTGTCCCTATGACCCAGCTAA	660
Db	601	AC	TTTGAGGAGATGAGAGTTGGAAGATGGGAAGGGCAAGTGTCCCTATGACCCAGCTAA	660
OY	661	GG	CCATGCTGGCTCTTCTGTGTGATGTAGCTGTACTTCGGCCACATCACTCAACTTCCTG	720
Db	661	GG	CCATGCTGGCTCTTCTGTGTGATGTAGCTGTACTTCGGCCACATCACTCAACTTCCTG	720
OY	721	GG	CAAGGAAACCATTAATCTCTGCGTACATGGGGCCCAACACTCTCATGAAAGACAGTAC	780
Db	721	GG	CAAGGAAACCATTAATCTCTGCGTACATGGGGCCCAACACTCTCATGAAAGACAGTAC	780
OY	781	CT	GCCTTTTGGCTCAACGAACTACACTTTTATGAGCTCTGCTATGTATACCTGAGAGTGT	840
Db	781	CT	GCCTCTTTTGGCTCAACGAACTACACTTTTATGAGCTCTGCTATGTATACCTGAGAGTGT	840
OY	841	GG	CAGCTTTCACGGGGGAGCAGCAGCAAGGCTACTTCTTTCAGGGAGCCGGCAGTGGAG	900
Db	841	GG	CAGCTTTCACGGGGGAGCAGCAGCAAGGCTACTTCTTTCAGGGAGCCGGCAGTGGAG	900
OY	901	TC	CCACTGTCTATGCGAGCAGTGTGTGCGCTGTGTGCGCTGTGTCTCAAGGGGATATG	960
Db	901	TC	CCACTGTCTATGCGAGCAGTGTGTGCGCTGTGTGCGCTGTGTCTCAAGGGGATATG	960
OY	961	GG	GGGCGCACGAGCCCTTCAGAGGAAGTGACACAGTCTCTGAAAGGGCGGCTGGCATGC	1020
Db	961	GG	GGGCGCACGAGCCCTTCAGAGGAAGTGACACAGTCTCTGAAAGGGCGGCTGGCATGC	1020
OY	1021	TC	TCTGCCCCGAACCTGGCAGCTCTACTTCAACCAAGCTGCAGCGGATGCACACCTTCAGAGAC	1080
Db	1021	TC	TCTGCCCCGAACCTGGCAGCTCTACTTCAACCAAGCTGCAGCGGATGCACACCTTCAGAGAC	1080
OY	1081	AC	CTCTGTGGCAACACACCACTTCTTTGGGTTTTTCAAGCAGAGTGGGGTGACATGTAC	1140
Db	1081	AC	CTCTGTGTGGCAACACACCACTTCTTTGGGTTTTTCAAGCAGAGTGGGGTGACATGTAC	1140
OY	1141	CT	GTGCGGCATCTGTGATACAGTTGGAAAGATTCAGCGGGGTGTTGAGGGCCCTAT	1200
Db	1141	CT	GTGCGGCATCTGTGATACAGTTGGAAAGATTCAGCGGGGTGTTGAGGGCCCTAT	1200
OY	1201	AA	GAGATACCATGAGGAAAGCCCAAGGGGAACCGCTACATGTGACCCCTGACCCAGCCT	1260
Db	1201	AA	GAGATACCATGAGGAAAGCCCAAGGGGAACCGCTACATGTGACCCCTGACCCAGCCT	1260
OY	1261	CG	GCCTGTGCTGTGCATTAACCACTGTGCATGGCGCCACGCTACACACAGCTCCCTGGAG	1320
Db	1261	CG	GCCTGTGCTGTGCATTAACCACTGTGCATGGCGCCACGCTACACACAGCTCCCTGGAG	1320
OY	1321	CT	ATCCCGAACATCTCTAATCTTGCTCAAGAACACCCGCTGATGTAGAGAGCAAGTGGGG	1380
Db	1321	CT	ATCCCGAACATCTCTAATCTTGCTCAAGAACACCCGCTGATGTAGAGAGCAAGTGGGG	1380
OY	1381	CC	TGCGTGTGAGACCCGCCCTCTGCTGTAAGAAAGGACCACTTCACCCACTGCTGGGCC	1440
Db	1381	CC	TGCGTGTGAGACCCGCCCTCTGCTGTAAGAAAGGACCACTTCACCCACTGCTGGGCC	1440
OY	1441	GAC	CGGGTTACAGGACTGTATGAGAGCCACTATACAGTGTCTGTTCATTTGGCACAGAGAC	1500
Db	1441	GAC	CGGGTTTACAGGACTGTATGAGAGCCACTATACAGTGTCTGTTCATTTGGCACAGAGAC	1500

QY	1501	GGCGGCGTCGACAAAGGCTGTAGACTGTGGGGCCCTGGGTTCACTCGATTGAGAGACTGCAG	1560
Db	1501	GGCTGGCTGTCAAGGCTGTGAGCCCTGGGGCCCTGGGTTCACTCGATTGAGAGACTGCAG	1560
QY	1561	CTGTTTGACAGAGAGCCCATGAGAAAGCCTGTGTGCTATTCTAGAGCAAGAAGCTGCTCTT	1620
Db	1561	CTGTTTGACAGAGAGCCCATGAGAAAGCCTGTGTGCTATTCTAGAGCAAAAAGCTGCTCTT	1620
QY	1621	GCCGGCTCCGCTCTCAGCTGTGTGACGTGCCCCCTGGCCGACTGCATTAAGTATGCTCC	1680
Db	1621	GCCGGCTCCGCTCTCAGCTGTGTGACGTGCCCCCTGGCCGACTGCATTAAGTATGCTCC	1680
QY	1681	TGTGAGACGTGTCTCTGCGCCCGGGAGCCCATGTAGCGGCTGAGACGTCAACACGACCGC	1740
Db	1681	TGTGAGACGTGTCTCTGCGCCCGGGAGCCCATGTAGCGGCTGAGACGTCAACACGACCGC	1740
QY	1741	TGTGTGGCGCTGGGTGGCCACTTTGGATTGTACTGTAGATCCAGCATGTGTATGATCCGAGAC	1800
Db	1741	TGTGTGGCGCTGGGTGGCGCACTCTGAGATCTCTACTGATCCAGCATGTGTATGATCCGAGAC	1800
QY	1801	ACTTCAGGCATTTTGCACACTCTCCGTGGCGATTAAGATACAGTATAGGCCCATCTTCCCAAAAAC	1860
Db	1801	ACTTCAGGCATCTGTCAACCTCTCCGTGGCGATTAAG -- AAGGTATAGGCCCACT - CCCAAAAAC	1857
QY	1861	ATCAGCGTGTGTGGCGGCGACAGACCTGTGTGCTGCGCTCCACACTCTCTCCAACTTGAGCC	1920
Db	1858	ATCAGCGTGTGTGGCGGCGACAGACCTGTGTGCTGCGCTCCACACTCTCTCTCCAACTTGAGCC	1917
QY	1921	CTG 1923	
Db	1918	CGG 1920	

[illegible]

Db 112 GCTCAGAGCGGGGGCTGCGCCATG6CCACACTG6GCTGTCTG6CTGTG6GCAAG 171
OY 204 GCTGTGGGCGCTGGGCAATGGGGCTGAGGTGTGTGGAACCTTGTGCCCTGAAGCACT 263
Db 172 GCTGTGGGCGCTGGGCAATGGGGCTGAGGTGTGTGGAACCTTGTGCCCTGAAGCACT 231
OY 264 GCTTCTGTGGGAGCTGGGCGACGAGTAGTGGGGGTTCTCCAGACCGGATCCAGACTT 323
Db 232 GCTTCTGTGGGAGCTGGGCGACGAGTAGTGGGGGTTCTCCAGACCGGATCCAGACTT 291
OY 334 CTTGACACTGACGCTGACGAGGAGCCCACTGGGCTTGTGTAGTGGGCGCCGAGAGGCTT 383
Db 232 CTTGACACTGACGCTGACGAGGAGCCCACTGGGCTTGTGTAGTGGGCGCCGAGAGGCTT 351
OY 384 GTTTCCCTTACAGATGAGAGGCGCTGTGAGCTGCAAGAGGCGATCTCTGGAGAGCCCTT 443
Db 352 GTTTCCTTACAGATGAGAGGCGCTGTGAGCTGCAAGAGGCGATCTCTGGAGAGCCCTT 411
OY 444 GGAGAAAGAACTGAGTGTATCCAGAAAGGAAAGAAACCAAGACCGAGTGTCTCAACTT 503
Db 412 GGAGAAAGAACTGAGTGTATCCAGAAAGGAAAGAAACCAAGACCGAGTGTCTCAACTT 471
OY 504 CATCCGCTTCTGACAGCCCTTACAAATGCTCCACCTGTGTGTGTGACCTTACGCTT 563
Db 472 CATCCGCTTCTGACAGCCCTTACAAATGCTCCACCTGTGTGTGTGACCTTACGCTT 531
OY 564 CCAGCCCAAGTGTACCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 623
Db 532 CCAGCCCAAGTGTACCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 591
OY 624 AGATGGAAAGGCAAGTGTATCCAGAAAGGAAAGAAACCAAGACCGAGTGTCTCAACTT 683
Db 592 AGATGGAAAGGCAAGTGTATCCAGAAAGGAAAGAAACCAAGACCGAGTGTCTCAACTT 651
OY 684 TGTGTAGCTGTACTGTGGGCGACACTCAACAATCTCTGGGCGAGGAAACCAATCTCTGGG 743
Db 652 TGTGTAGCTGTACTGTGGGCGACACTCAACAATCTCTGGGCGAGGAAACCAATCTCTGGG 711
OY 744 TAACTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 803
Db 712 TAACTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 771
OY 804 TCACCTTGTAGGCTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 863
Db 772 TCACCTTGTAGGCTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 831
OY 864 CAAGGTCTACTTCTTCTCAAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 923
Db 832 CAAGGTCTACTTCTTCTCAAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 891
OY 924 GCTGCTGTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 983
Db 892 GCTGCTGTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 951
OY 984 GAAAGTGACACAGTGTCTGAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1043
Db 952 GAAAGTGACACAGTGTCTGAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1011
OY 1044 CTTCAACAGCTGACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1103
Db 1012 CTTCAACAGCTGACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1071
OY 1104 CTTTGGGGTTTTTCAAGCAGCACTGTGTGACATGTGTGTGTGTGTGTGTGTGTGTGT 1163
Db 1072 CTTTGGGGTTTTTCAAGCAGCACTGTGTGACATGTGTGTGTGTGTGTGTGTGTGTGT 1131
OY 1164 GTTGAAGAGATCCAGGGGTTTGTAGGGCCCTTAAAGAGTACATGTGTGTGTGTGTGTGT 1223
Db 1132 GTTGAAGAGATCCAGGGGTTTGTAGGGCCCTTAAAGAGTACATGTGTGTGTGTGTGTGT 1191
OY 1224 GAAAGTGACACAGTGTCTGAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1283
Db 1192 GAAAGTGACACAGTGTCTGAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1251

OY 1284 CTGECATGCGCGCAGCGGCTTACACCAAGCTCCCTGAGCTTACCCGACACATCTCAACTT 1343
Db 1252 CTGECATGCGCGCAGCGGCTTACACCAAGCTCCCTGAGCTTACCCGACACATCTCAACTT 1311
OY 1344 CGTCAAGAAAGACACCGGCTATGAGAGAGAGAGTGGGGGCTCGGTGAGAGCGCCCTGCT 1403
Db 1312 CGTCAAGAAAGACACCGGCTATGAGAGAGAGAGTGGGGGCTCGGTGAGAGCGCCCTGCT 1371
OY 1404 CGTGAAGAAAGGACACCAACTTACCCAGCTGTGTGGCGGAGCGGCTTACAGAGCTTATGAG 1463
Db 1372 CGTGAAGAAAGGACACCAACTTACCCAGCTGTGTGGCGGAGCGGCTTACAGAGCTTATGAG 1431
OY 1464 AGCCACTATACAGT 1523
Db 1432 AGCCACTATACAGT 1491
OY 1524 CTTGGGGGCGGCTGACCTGATTTGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1583
Db 1492 CTTGGGGGCGGCTGACCTGATTTGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1551
OY 1584 AAGCCTGT 1643
Db 1552 AAGCCTGT 1611
OY 1644 GCAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1703
Db 1612 GCAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1671
OY 1704 GGAGCCCTTATGCGGCTGTGAGGCGTCAACACAGCGGCTGTGTGTGTGTGTGTGTGTGT 1763
Db 1672 GGAGCCCTTATGCGGCTGTGAGGCGTCAACACAGCGGCTGTGTGTGTGTGTGTGTGTGT 1731
OY 1764 TGGATCTTACTGATTCACAGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1823
Db 1732 TGGATCTTACTGATTCACAGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1791
OY 1824 TGGCAATTAAGATACAGT 1883
Db 1792 TGGCAATTAAGATACAGT 1848
OY 1884 CTTGT 1942
Db 1848 CTTGT 1907

RESULT 5
AX174731 3293 bp DNA linear PAT 03-JUL-2001
LOCUS AX174731
DEFINITION Sequence 25 from Patent WO0142285.
ACCESSION AX174731
VERSION AX174731.1 GI:14598254
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Hom sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 Yue,H., Tang,Y.T., Lal,P., Burford,N., Azimzal,Y., Patterson,C.,
Baughn,M.R., Lu,D.A., Shah,P. and Au-Young,J.
Extracellular matrix and cell adhesion proteins as well as genes
encoding them
Patent: WO 0142285-A 25 14-JUN-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1..3293
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 1674661CB1"
BASE COUNT 599 a 978 c 1052 g 664 t
ORIGIN

Query Match	77.18	Score 1761	DB 6	Length 3293
Best Local Similarity	97.48	Pred. No. 0		
Matches 1811	Conservative 0	Mismatches 45	Indels 3	Gaps 2
QY	84	CAGGGCGGACCGGACCTCCGCTCCAGCGCGCCGACACAGACGACCTGGCCGAA	143	
DB	52	CCGAAACGCGAGGGCGCGCGCGGAGCTGGCGCAATAGGACACCTGGCCGAA	111	
QY	144	GCTCAGACGCGGGCGCTGGCCATGGCCCACTGGGCTGTGTGGCTGTGGCAGAG	203	
DB	112	GCTCAGACGCGGGCGCTGGCCATGGCCCACTGGGCTGTGTGGCTGTGGCAGAG	171	
QY	204	GCTGTGGGCGCTGGGCTATGGGCTGAGGTGTGGACCTGTGTGGCTGTAAGACGT	263	
DB	172	GCTGTGGGCGCTGGGCTATGGGCTGAGGTGTGGACCTGTGTGGCTGTAAGACGT	231	
QY	264	GTTCTTGGGGAGCTGGGCGCGGTATGCGGGGTCTCCAGACCGGCAATCCAGACT	323	
DB	232	GTTCTTGGGGAGCTGGGCGCGGTATGCGGGGTCTCCAGACCGGCAATCCAGACT	291	
QY	324	CCTGACACTGACGCTGACGAGGCCACTGGGCTTCTGTACGTGGGCGCCGAGAGGCT	383	
DB	292	CCTGACACTGACGCTGACGAGGCCACTGGGCTTCTGTACGTGGGCGCCGAGAGGCT	351	
QY	384	GTTTGCTTACGATGAGGCGCTGGAGCTGCAAGAGCGATCTCTGGGAGGCCCGCT	443	
DB	352	GTTTGCTTACGATGAGGCGCTGGAGCTGCAAGAGCGATCTCTGGGAGGCCCGCT	411	
QY	444	GGAAGAAAGACGAGTGTATCCAGAAAGGAAAGAACACAGACGAGTGTCAACT	503	
DB	412	GGAAGAAAGACGAGTGTATCCAGAAAGGAAAGAACACAGACGAGTGTCAACT	471	
QY	504	CATCCCTTCCCTGACGCGCTACAAATGCTCCACCTGTACGTGTGGGCACTACGCTT	563	
DB	472	CATCCCTTCCCTGACGCGCTACAAATGCTCCACCTGTACGTGTGGGCACTACGCTT	531	
QY	564	CCAGCCCAAGTGCACCTACGTCACATGCTCACCTTCTGTGAGTGAAGAGTGA	623	
DB	532	CCAGCCCAAGTGCACCTACGTCACATGCTCACCTTCTGTGAGTGAAGAGTGA	591	
QY	624	AGATGGGAAGGGAAGTGTCCCTATGACCCAGTAAAGGCGATGCTGTCTTGTGA	683	
DB	592	AGATGGGAAGGGAAGTGTCCCTATGACCCAGTAAAGGCGATGCTGTCTTGTGA	651	
QY	684	TGCTGAGCTGTACTCGGCGCACACTCAACAATCTCTGGGCGACGAAACCACTTACCTGCG	743	
DB	652	TGCTGAGCTGTACTCGGCGCACACTCAACAATCTCTGGGCGACGAAACCACTTACCTGCG	711	
QY	744	TAAACATGGGGCGCCACCACTCCATGAAGACAGATACCTGCTTGTGGCTCAACGAAAC	803	
DB	712	TAAACATGGGGCGCCACCACTCCATGAAGACAGATACCTGCTTGTGGCTCAACGAAAC	771	
QY	804	TCACTTGTAGGCTGTGCGCTATGTAAGTGAAGTGTGGGCGAGTCAACGGGGGACGACGA	863	
DB	772	TCACTTGTAGGCTGTGCGCTATGTAAGTGAAGTGTGGGCGAGTCAACGGGGGACGACGA	831	
QY	864	CAAGGTCTACTTCTTCTAGGAGCGGCGAGTGAATCCGATGCTATCCGAGCAGGT	923	
DB	832	CAAGGTCTACTTCTTCTAGGAGCGGCGAGTGAATCCGATGCTATCCGAGCAGGT	891	
QY	924	GCTGCTGTGTGGCGCGCTGTCTGCAAGGCGATGTGGGGCGCGACGACCTGTGAGAG	983	
DB	892	GCTGCTGTGTGGCGCGCTGTCTGCAAGGCGATGTGGGGCGCGACGACCTGTGAGAG	951	
QY	984	GAAGTGACACCTCTCTGAAGGCGGGCGTGGCATGTCTGCCCCGAACTGGGAGCTCTA	1043	
DB	952	GAAGTGACACCTCTCTGAAGGCGGGCGTGGCATGTCTGCCCCGAACTGGGAGCTCTA	1011	
QY	1044	CTTCAACACAGCTGCAAGGCGATGACACCTGCGAGACCTCTGGGCAACACACCTT	1103	
DB	1012	CTTCAACACAGCTGCAAGGCGATGACACCTGCGAGACCTCTGGGCAACACACCTT	1071	
QY	1104	CTTTGGGGTTTTTCAAGACAGTGGGGTGAACATGTACTGTGGGCAATCTGTAGTAACA	1163	

DB	1072	CTTTGGGGTTTTTCAAGACAGTGGGGTGAACATGTACTCTGGCCATCTGTGAGTAACA	1131	
QY	1164	GTTGGAAGAGATACAGCGGGTCTTTTGAGGGCCCTTAAAGAGTACCATATGAGAGCCCA	1223	
DB	1132	GTTGGAAGAGATACAGCGGGTCTTTTGAGGGCCCTTAAAGAGTACCATATGAGAGCCCA	1191	
QY	1224	GAAGTGGAGCCGCTACACTGACCTGTACCAGCCCTGGGCTGTGCTGTGATTAAACA	1283	
DB	1192	GAAGTGGAGCCGCTACACTGACCTGTACCAGCCCTGGGCTGTGCTGTGATTAAACA	1251	
QY	1284	CTGGCATGGCGCCACGCTACACACAGCTCCCTGGAGCTAACCCGAAACATCTCAACT	1343	
DB	1252	CTGGCATGGCGCCACGCTACACACAGCTCCCTGGAGCTAACCCGAAACATCTCAACT	1311	
QY	1344	CGTCAAGAACACCCGCTGATGAGAGAGACAGTGGGGCTCGGTGAGCGCCCGCTGCT	1403	
DB	1312	CGTCAAGAACACCCGCTGATGAGAGAGACAGTGGGGCTCGGTGAGCGCCCGCTGCT	1371	
QY	1404	CGTGAAGAAAGGCGCACAACTTACACCCACTGTGGCGCGACCGGGTTACAGGACTGTAG	1463	
DB	1372	CGTGAAGAAAGGCGCACAACTTACACCCACTGTGTGGCGCGACCGGGTTACAGGACTGTAG	1431	
QY	1464	AGCCACCTATACAGTGTCTTCAATGGCGACAGAGAGCGGTGCTGTCAAGGCTGTAG	1523	
DB	1432	AGCCACCTATACAGTGTCTTCAATGGCGACAGAGAGCGGTGCTGTCAAGGCTGTAG	1491	
QY	1524	CTTGGGGCCCTGGGTTACCTGATTTGAGAGCTGACCTGTGTTGACAGAGAGCCATGAG	1583	
DB	1492	CTTGGGGCCCTGGGTTACCTGATTTGAGAGCTGACCTGTGTTGACAGAGAGCCATGAG	1551	
QY	1584	AAGCTGTGTCTATCTCAGAGCAAGAGCTGCTTGTGGCGGCTCCGCTGTACGCTGCT	1643	
DB	1552	AAGCTGTGTCTATCTCAGAGCAAGAGCTGCTTGTGGCGGCTCCGCTGTGTACGCTGCT	1611	
QY	1644	GCAAGCTGCCGCTGGCGCACTGCAATMAATGATGCTGCTGTGACAGACTGTCTCGCGCG	1703	
DB	1612	GCAAGCTGCCGCTGGCGCACTGCAATMAATGATGCTGCTGTGACAGACTGTCTCGCGCG	1671	
QY	1704	GGACCCCTATTTGGGCGCTGAGAGCGTCAACACAGCGCTGTGTGGCGCTGGGTCGCACTT	1763	
DB	1672	GGACCCCTATTTGGGCGCTGAGAGCGTCAACACAGCGCTGTGTGGCGCTGGGTCGCACTT	1731	
QY	1764	TGATCTTTACTGATTCAGAGATGTGATGACCTGGGCACTTACAGATTGCAACTCCG	1823	
DB	1732	TGATCTTTACTGATTCAGAGATGTGATGACCTGGGCACTTACAGATTGCAACTCCG	1791	
QY	1824	TGGCAGTAAATACAGTCAGGCGCCACTTCCAAAACATCAGGGTGTGGCGGACAGA	1883	
DB	1792	TGGCAGTAAAG--TAAAGTCAGGCGCCACT--CCCAAAAACATCAGGGTGTGGCGGACAGA	1848	
QY	1884	CTGTGGTGTGCGCGCGACACTTCTCCCAACTTGTGGCCCTGCCGACCTCAACCCCGAGG	1942	
DB	1849	CTGTGGTGTGCGCGCGACACTTCTCCCAACTTGTGGCCCATGCCGCTGTGAGCTTTGGGG	1907	

RESULT 6	AX528275	3868 bp	DNA	linear	PAT 21-NOV-2002
LOCUS	AX528275				
DEFINITION	Sequence 13 from Patent WO206339.				
ACCESSION	AX528275				
VERSION	AX528275.1	GI:25172575			
KEYWORDS					
SOURCE					
ORGANISM					
	Homo sapiens (human)				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
REFERENCE					
AUTHORS	Spaderna,S.K., Tchernev,V., Liu,X., Shenoy,S., Spytek,R., Zernhusen,B., Paturusan,M., Taupier,R.J., Kastelli,L., Grosse,M.M., Szekeres,E.S., Alsbrook,J.T., Lepley,D.M., Shen,L., Burgess,C.E., Shimkets,R.A. and Padigaru,M.				
TITLE	Proteins and nucleic acids encoding same				

JOURNAL Patent: WO 0206339-A 13 24-JAN-2002;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source 1. 3868
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 717 a 1106 c 1207 g 838 t
ORIGIN
Query Match 76.1%; Score 1738; DB 6; Length 3868;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1782; Conservative 0; Mismatches 35; Indels 3; Gaps 2;
123 CAGCAGACAGCTGGCCCTGAGAGCTCAGAGCCGGGCGCTGCGCCATGAGCCCACTAGGCGC 182
1410 CAGCAGACAGCTGGCCCTGAGAGCTCAGAGCCGGGCGCTGCGCCATGAGCCCACTAGGCGC 469
183 TGTCTGCTGCTGCTGAGAGAGCTGTGGGCGCTGGGCGCTGAGGCTGAGGCTGTGGTGA 242
470 TGTCTGCTGCTGCTGAGAGAGCTGTGGGCGCTGGGCGCTGAGGCTGAGGCTGTGGTGA 529
243 CCTTGTGCGCGGTAAAGAGCTGTCTTCTGAGGAGCTGCGCAGGCTAGTACGGCGGTTCTC 302
530 CCTTGTGCGCGGTAAAGAGCTGTCTTCTGAGGAGCTGCGCAGGCTAGTACGGCGGTTCTC 589
303 CCAAGCCGGCATTCAGAGACTTCTGACACTAGCTGACGAGAGCCCACTGGGCTTCTGTA 362
590 CCAAGCCGGCATTCAGAGACTTCTGACACTAGCTGACGAGAGCCCACTGGGCTTCTGTA 649
363 CGTGGGCGCGGAGAGGCGCTGTTGCGCTTACAGATGAGAGGCGCTGAGCTGCGAAGAGC 422
650 CGTGGGCGCGGAGAGGCGCTGTTGCGCTTACAGATGAGAGGCTGAGCTGCGAAGAGC 709
423 GATCTCTGAGAGGCGCGCTGAGAGAGAGAGAGCTAGTATCCAGAAAGGAGAGACAA 482
710 GATCTCTGAGAGGCGCGCTGAGAGAGAGAGAGCTAGTATCCAGAAAGGAGAGACAA 769
483 CCAAGCCGAGTCTTCAACTTCACTCCGCTTCTGAGCCCTACAACTGCTCCCACTGTA 542
770 CCAAGCCGAGTCTTCAACTTCACTCCGCTTCTGAGCCCTACAACTGCTCCCACTGTA 829
543 CGTGTGAGCACTAGCGCTTCCAGCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 602
830 CGTGTGAGCACTAGCGCTTCCAGCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 889
603 TTTGAGAGCAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 662
890 TTTGAGAGCAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 949
663 CCAATGCTGAGCTTCTTGTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 722
950 CCAATGCTGAGCTTCTTGTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1009
723 CAGGAGACCCATTATCCGCTGATACATGAGGCGCCACACAGTCCATGAGAGAGAGTACT 782
1010 CAGGAGACCCATTATCCGCTGATACATGAGGCGCCACACAGTCCATGAGAGAGAGTACT 1069
783 GGCCTTTTGGCTCAACGAACTCACTTTGAGGCTGCTGCTATGATGATGATGATGATGATG 842
1070 GGCCTTTTGGCTCAACGAACTCACTTTGAGGCTGCTGCTATGATGATGATGATGATGATG 1129
843 CAGCTTCAAGGGGAGCAGCAGCAAGTCTACTTCTTCAAGGAGCGGGAGAGTGAATC 902
1130 CAGCTTCAAGGGGAGCAGCAGCAAGTCTACTTCTTCAAGGAGCGGGAGAGTGAATC 1189
903 CGACTGCTATGCGCAGCAGTGTGGCTGCTGTGGCGCGTGTCTGCAAGGGCGATATGG 962
1190 CGCTGCTATGCGCAGCAGTGTGGCTGCTGTGGCGCGTGTCTGCAAGGGCGATATGG 1249
963 GGGCGCAGGAGCCTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1022
1250 GGGCGCAGGAGCCTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1309

1033 TGCCCCGAACAGGAGCTCTACCTTCACAGCTGAGGCGATGCACACCTGTCAGAGAC 1082
1310 TGCCCCGAACAGGAGCTCTACCTTCACAGCTGAGGCGATGCACACCTGTCAGAGAC 1369
1083 CTCCTGGCACAACACACCTTCTTGGGTTTTTAACAGCAGTGGGGTGAATATCT 1142
1370 CTCCTGGCACAACACACCTTCTTGGGTTTTTAACAGCAGTGGGGTGAATATCT 1429
1143 GTCGCGCATCTGTAGTACAGAGTGAAGAGATCCAGGGGGTTTGAAGGGCCCTATTA 1202
1430 GTCGCGCATCTGTAGTACAGAGTGAAGAGATCCAGGGGGTTTGAAGGGCCCTATTA 1489
1203 GAGATACCATGAGAGAGCCAGAAAGTGGGACCGCTACACTGACCTGTACCCAGCCCTC 1262
1490 GAGATACCATGAGAGAGCCAGAAAGTGGGACCGCTACACTGACCTGTACCCAGCCCTC 1549
1263 GCTTGGCTGCTGCTATTAACAACTGGCATTCGGGCGCCAGCGCTACACACCTGCTGAGCT 1322
1550 GCTTGGCTGCTGCTATTAACAACTGGCATTCGGGCGCCAGCGCTACACACCTGCTGAGCT 1609
1333 ACCGACAAACATCTCAACTTCGTCAAGAGACCCGCTGATGAGAGAGAGTGGGCG 1382
1610 ACCGACAAACATCTCAACTTCGTCAAGAGACCCGCTGATGAGAGAGAGTGGGCG 1669
1383 TCGGTGAGAGCCGCCCTGCTGCTGAGAGAGGAGACCAACTTACCCACCTGCTGGCGA 1442
1670 TCGGTGAGAGCCGCCCTGCTGCTGAGAGAGGAGACCAACTTACCCACCTGCTGGCGA 1729
1443 CCGGTTACAGAGCTGATGAGAGAGCCCACTATACATGCTGTTCAATGGCAGAGAGAG 1502
1730 CCGGTTACAGAGCTGATGAGAGAGCCCACTATACATGCTGTTCAATGGCAGAGAGAG 1789
1503 CTGCTGCTCAAGGCTGTGAGGCTGGGCGCTGAGGCTGATGAGAGAGTCACT 1562
1790 ATGCTGCTCAAGGCTGTGAGGCTGGGCGCTGAGGCTGATGAGAGAGTCACT 1849
1563 GTTTCACAGAGAGCCCACTGAGAGAGCTGCTGCTATCTCAGAGCAAGAGTCTTCTTGC 1622
1850 GTTTCACAGAGAGCCCACTGAGAGAGCTGCTGCTATCTCAGTGCAGAGAGTCTTCTTGC 1909
1623 CGGCTCCGCGCTCAGCTGCTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1682
1910 CGGCTCCGCGCTCAGCTGCTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1969
1683 TGCAGACTGTGCTCGCGCCGAGCCCTATGCTGCTGAGAGAGCTCAACACAGCGCTG 1742
1970 TGCAGACTGTGCTCGCGCCGAGCCCTATGCTGCTGAGAGAGCTCAACACAGCGCTG 2029
1743 TGTGGCGTGGGTGGCACTTTGATCTTACTGATCCAGCATGTGATGATGATGATGATG 1802
2030 TGTGGCGTGGGTGGCACTTTGATCTTACTGATCCAGCATGTGATGATGATGATGATG 2089
1803 TTTACAGCATTTGCAACCTCCCGTGAAGTAAAGTCAAGGAGCCCACTGCAAGAAAT 1862
2090 TTTACAGCATTTGCAACCTCCCGTGAAGTAAAGTCAAGGAGCCCACTGCAAGAAAT 2146
1863 CAGGCTGCTGAGGAGCAGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1922
2147 CAGGCTGCTGAGGAGCAGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2206
1923 GCCGACTCCAAACCCGAGG 1942
2207 TGCCCGCTGACCTTTGGG 2226
RESULT 7
AX528279 3112 bp DNA linear PAT 21-NOV-2002
LOCUS AX528279
DEFINITION Sequence 17 from Patent WO0206339.
ACCESSION AX528279
VERSION AX528279.1 GI:25172577
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS Spaderna, S.K., Tcherney, V., Liu, X., Shenoy, S., Spytek, R.,
Zerhisen, B., Palturalian, M., Taupier, R.J., Rastelli, L., Grosse, W.M.,
Szikeres, E.S., Alsobrook, J.I., Lepley, D.M., Shen, L., Burgess, C.E.,
Shimkets, R.A. and Padigaru, M.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 0206339-A 17 24-JAN-2002;
Curagen Corporation (US)
FEATURES
source 1. .3112
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 559 a 948 c 957 g 648 t
ORIGIN
Query Match 75.4%; Score 1722; DB 6; Length 3112;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 1772; Conservative 0; Mismatches 45; Indels 3; Gaps 2;
QY 123 CAGCAGACAGCTGCGCTGAAGCTCAGAGCCGGGGCGTGGCCATGGCCACACTGGGC 182
DB 61 CAGCAGACAGCTGCGCTGAAGCTCAGAGCCGGGGCGTGGCCATGGCCACACTGGGC 120
QY 183 TGTCTGGCTGCTGCGCAGCAAGGCTGCGGGCTGGGCTATGGGGCTAGGTGTGGAA 242
DB 121 TGTCTGGCTGCTGCGCAGCAAGGCTGCGGGCTGGGCTATGGGGCTAGGTGTGGAA 180
QY 243 CTTTGTGCGCGTGAAGACAGTCTTCTGGGAGCTGGCCAGCGTAGTACGGCGTCTC 302
DB 181 CTTTGTGCGCGTGAAGACAGTCTTCTGGGAGCTGGCCAGCGTAGTACGGCGTCTC 240
QY 303 CCGACCGCGATCCAGACTTCTGACACTGACGCTGACGGAGCCACCTGGCTTCTGA 362
DB 241 CCGACCGCGATCCAGACTTCTGACACTGACGCTGACGGAGCCACCTGGCTTCTGA 300
QY 363 CGTGGGGCGCCGAGAGGCCCTGTTGCTCAGACTGGAAGCCCTGAGCTGCAAGAGC 422
DB 301 CGTGGGGCGCCGAGAGGCCCTGTTGCTCAGACTGGAAGCCCTGAGCTGCAAGAGC 360
QY 423 GATCTCTGGAGAGCCCGCTGGAGAAAGACTGATGATCCAGAAAGGAGAACAA 482
DB 361 GATCTCTGGAGAGCCCGCTGGAGAAAGACTGATGATCCAGAAAGGAGAACAA 420
QY 483 CCGACCGAGTGTCTCAACTTCATCGCTTCTGACAGCCCTACAGCCCTCCACTGTA 542
DB 421 CCGACCGAGTGTCTCAACTTCATCGCTTCTGACAGCCCTACAGCCCTCCACTGTA 480
QY 543 CGCTGTGGACCTAGAGCCCTTCCAGCCAAAGTCACTAGTCAACATGCTTACCTTAC 602
DB 481 CGCTGTGGACCTAGAGCCCTTCCAGCCAAAGTCACTAGTCAACATGCTTACCTTAC 540
QY 603 TTTGGAGCATGAGAGTTTGAAGATGGAAGGGCAAGTGTCCCTATGACCCAGCTAAGG 662
DB 541 TTTGGAGCATGAGAGTTTGAAGATGGAAGGGCAAGTGTCCCTATGACCCAGCTAAGG 600
QY 663 CCACTGTGCGCTTCTGTGATGTGAGCTGTAAGTGGCCACACTCAACACTTCTGGG 722
DB 601 CCACTGTGCGCTTCTGTGATGTGAGCTGTAAGTGGCCACACTCAACACTTCTGGG 660
QY 723 CAGGGAACCATATATCTGCTGATGAAGTGGGGCCCACTCACTGAAGACAGATACCT 782
DB 661 CAGGGAACCATATATCTGCTGATGAAGTGGGGCCCACTCACTGAAGACAGATACCT 720
QY 783 GGGCTTTTGGCTCAAGAACCTCACTTGTAGGCTGTGCTATGTAAGTGAAGTGGG 842
DB 721 GGGCTTTTGGCTCAAGAACCTCACTTGTAGGCTGTGCTATGTAAGTGAAGTGGG 780
QY 843 CAGCTTCAAGGGGAGCAGCAAGGTCTACTTCTTCAAGGAGCGGGCACTGAGTCT 902

DB 781 CAGCTTCAAGGGGAGCAGCAAGGTCTACTTCTTCAAGGAGCGGGCAGTGAAGTCT 840
QY 903 CGACTGTATGCGAGAGAGTGTGCTCTGTGGGCCCGTGTGCAAGGGCCATATAGG 962
DB 841 CGACTGTATGCGAGAGAGTGTGCTCTGTGGGCCCGTGTGCAAGGGCCATATAGG 900
QY 963 GGGCGCAGGACCCCTGCGAGAGAAAGTGGACACCTTCTGTAAGGCGGGCTGCACTCTC 1022
DB 901 GGGCGCAGGACCCCTGCGAGAGAAAGTGGACACCTTCTGTAAGGCGGGCTGCACTCTC 960
QY 1023 TGCCTGCAAGTGGCTGCTACTTCAACACAGCTGCAAGGCGATGCAACCTGCAAGAC 1082
DB 961 TGCCTGCAAGTGGCTGCTACTTCAACACAGCTGCAAGGCGATGCAACCTGCAAGAC 1020
QY 1083 CTCTGCGCAAGACCACTTCTTGGGGTTTTTCAAGCACAGTGGGGGTGACATGTACT 1142
DB 1021 CTCTGCGCAAGACCACTTCTTGGGGTTTTTCAAGCACAGTGGGGGTGACATGTACT 1080
QY 1143 GTGCGCATCTGTAGTACAGTTGGAAGAGATCCAGGGGTGTTTGAAGGCGCCCTATA 1202
DB 1081 GTGCGCATCTGTAGTACAGTTGGAAGAGATCCAGGGGTGTTTGAAGGCGCCCTATA 1140
QY 1203 GGAATACCATGAGAACCCGAGAAAGTGGACCGCTACACTGACCTGTACCAAGCCCTCG 1262
DB 1141 GGAATACCATGAGAACCCGAGAAAGTGGACCGCTACACTGACCTGTACCAAGCCCTCG 1200
QY 1263 GCCGTGCTGTGCTTAAACAATGCGCATGCGCGCCACAGGCTACACAGCTCCTGTAGCT 1332
DB 1201 GCCGTGCTGTGCTTAAACAATGCGCATGCGCGCCACAGGCTACACAGCTCCTGTAGCT 1260
QY 1323 ACCGACAAATCTCTCAACTTCTGCAAGAACACCCGCTGATGAGAGAGCAGTGGGGCC 1382
DB 1261 ACCGACAAATCTCTCAACTTCTGCAAGAACACCCGCTGATGAGAGAGCAGTGGGGCC 1320
QY 1383 TCGGTGAGCGCGCCCTGCTCTGTAAGAAAGGACCAACCTTCAACCACTGTGTGCGGA 1442
DB 1321 TCGGTGAGCGCGCCCTGCTCTGTAAGAAAGGACCAACCTTCAACCACTGTGTGCGGA 1380
QY 1443 CCGGCTTACAGGACTTATGAGAGCCACTTATACAGTCTGTTCATTTGGGCAAGAGACGG 1502
DB 1381 CCGGCTTACAGGACTTATGAGAGCCACTTATACAGTCTGTTCATTTGGGCAAGAGACGG 1440
QY 1503 CTGGCTGCTCAAGGCTGTGAGCTGGGGCCCTGGGTTCACTGTATGAGAGCTGACGCT 1562
DB 1441 CTGGCTGCTCAAGGCTGTGAGCTGGGGCCCTGGGTTCACTGTATGAGAGCTGACGCT 1500
QY 1563 GTTTGACAGAGCCCATGAGAAAGCTGTGTCTATCTCAGAGCAAGAAAGCTGCTTTGC 1622
DB 1501 GTTTGACAGAGCCCATGAGAAAGCTGTGTCTATCTCAGAGCAAGAAAGCTGCTTTGC 1560
QY 1623 CGGCTCCGCTCTCAGTGTGAGCGAGTGGCCCGTGGCGAGTGCATTAAGTATGCTCTCTG 1682
DB 1561 CGGCTCCGCTCTCAGTGTGAGCGAGTGGCCCGTGGCGAGTGCATTAAGTATGCTCTCTG 1620
QY 1683 TGCAGACTGTGTCTGCGCCGCGGACCCCTATTGGCGCTGGAGCGTGAACACAGCGCTG 1742
DB 1621 TGCAGACTGTGTCTGCGCCGCGGACCCCTATTGGCGCTGGAGCGTGAACACAGCGCTG 1680
QY 1743 TGTGGCCGTGGGTGGCACTTTGGATCTTACTGATCCAGCANTGTATGACTGTGGACAC 1802
DB 1681 TGTGGCCGTGGGTGGCACTTGTGATCTCTACTGATCCAGCANTGTATGACTGTGGACAC 1740
QY 1803 TTTAGGCAATTTGCAACTCGGTGGCACTAAGATACATCAGGCGCCACTTCCCAAAAACAT 1862
DB 1741 TTTAGGCAATTTGCAACTCGGTGGCACTAAG--AAATTCAGGGCCCACT-CCCAAAAACAT 1797
QY 1863 CAGGTGTGTGGCGGCAAGACCTGTGTGCTGCGCTGCGACCTCTCTCAACTGTGGCGCT 1922
DB 1798 CAGGTGTGTGGCGGCAAGACCTGTGTGCTGCGCTGCGACCTCTCTCAACTGTGGCGCA 1857
QY 1923 GCCCGACTCCAAACCCGAGG 1942
II II I I I I I

DB 1858 TGCCCGCTGACCTTGGGG 1877

RESULT 8
AX528277
LOCUS AX528277
DEFINITION Sequence 15 from Patent WO0206339.
ACCESSION AX528277
VERSION AX528277.1 GI:25172576
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Spaderne, S.K., Therneau, V., Liu, X., Shenoy, S., Spytek, K.,
Zehresen, B., Paturajan, M., Taupier, R.J., Raselli, L., Grose, W.M.,
Szekeres, E.S., Alsobrook, J.I., Lepley, D.M., Shen, L., Burgess, C.E.,
Shinkens, R.A., and Padigaru, M.
Proteins and nucleic acids encoding same
Patent: WO 0206339-A 15 24-JAN-2002;
Curagen Corporation (US)
TITLE
JOURNAL Location/Qualifiers
FEATURES
SOURCE 1..2558
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 468 a 788 c 801 g 501 t
ORIGIN

Query Match 74.4%; Score 1699; DB 6; Length 2558;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1749; Conservative 0; Mismatches 45; Indels 3; Gaps 2;

146 TCAGAGCCGCGGCGCTGCGCCATGCGCCCAACACTGCGCTGCTGCGCAGCAAGGC 205
DB 1 TCAGAGCCGCGGCGCTGCGCCATGCGCCCAACACTGCGCTGCTGCGCAGCAAGGC 60

206 TGTGGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGG 265
DB 1 TGTGGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGG 120

61 TGTGGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGG 120

266 CTTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGG 325
DB 1 CTTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGG 180

121 CTTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGG 180

326 TGACACTGACCTGACGCGAGCGCCACTGCGCTTCTGTACGTGGCGCGCGAGGCGCTGT 385
DB 1 TGACACTGACCTGACGCGAGCGCCACTGCGCTTCTGTACGTGGCGCGCGAGGCGCTGT 240

181 TGACACTGACCTGACGCGAGCGCCACTGCGCTTCTGTACGTGGCGCGCGAGGCGCTGT 240

386 TTGCGCTTACGATGAGGCGCTTGGAGCTGCAAGAGCGCATCTCTGGAGGCGCGCGCTGT 445
DB 1 TTGCGCTTACGATGAGGCGCTTGGAGCTGCAAGAGCGCATCTCTGGAGGCGCGCGCTGT 300

241 CTGCACTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300

446 AGAAGAAGACGTGAGTGTATCCAGAAAGGAGAAACACACAGAGCGAGTTCACACTTCA 505
DB 1 AGAAGAAGACGTGAGTGTATCCAGAAAGGAGAAACACACAGAGCGAGTTCACACTTCA 360

301 AGAAGAAGACGTGAGTGTATCCAGAAAGGAGAAACACACAGAGCGAGTTCACACTTCA 360

506 TCCGCTTCTGCGAGCGCTGCAATGCGCTCCACCTGTACGTGTGGGCGACTACGCGCTTCC 565
DB 1 TCCGCTTCTGCGAGCGCTGCAATGCGCTCCACCTGTACGTGTGGGCGACTACGCGCTTCC 420

361 TCCGCTTCTGCGAGCGCTGCAATGCGCTCCACCTGTACGTGTGGGCGACTACGCGCTTCC 420

566 AGCCCAAGTGCACCTACGTCAACATGCTACCTTCTGAGAGTGAAG 625
DB 1 AGCCCAAGTGCACCTACGTCAACATGCTACCTTCTGAGAGTGAAG 480

421 AGCCCAAGTGCACCTACGTCAACATGCTACCTTCTGAGAGTGAAG 480

626 ATGGGAAGGGAAGTGTCCCTATGACCCAGTAAGGCGCATGCTGCGCTTCTGTGGATG 685
DB 1 ATGGGAAGGGAAGTGTCCCTATGACCCAGTAAGGCGCATGCTGCGCTTCTGTGGATG 540

481 ATGGGAAGGGAAGTGTCCCTATGACCCAGTAAGGCGCATGCTGCGCTTCTGTGGATG 540

686 GTGAGCTGTACTGCGGCGCACTCAACAACTTCTGCGGCGAGGAAGCCATTATCCGCGTA 745
DB 1 GTGAGCTGTACTGCGGCGCACTCAACAACTTCTGCGGCGAGGAAGCCATTATCCGCGTA 600

746 ACATGGGCGCCCGACCTCCATGAAAGACAGACTACCTGGCCCTTTGGCTCAACGAACCTTC 805
DB 1 ACATGGGCGCCCGACCTCCATGAAAGACAGACTACCTGGCCCTTTGGCTCAACGAACCTTC 660

601 ACATGGGCGCCCGACCTCCATGAAAGACAGACTACCTGGCCCTTTGGCTCAACGAACCTTC 660

806 ACTTTGTAGGCTCTGCTGCTATGTAAGAGTGTGGGCGAGCTTCAACGGGGGAGCGAGCA 865
DB 1 ACTTTGTAGGCTCTGCTGCTATGTAAGAGTGTGGGCGAGCTTCAACGGGGGAGCGAGCA 720

661 ACTTTGTAGGCTCTGCTGCTATGTAAGAGTGTGGGCGAGCTTCAACGGGGGAGCGAGCA 720

866 AGCTCTACTTCTTCTGAGGAGCGGCGAGTGGAGTCCGACTGCTATGGCGAGCGAGTGG 925
DB 1 AGCTCTACTTCTTCTGAGGAGCGGCGAGTGGAGTCCGACTGCTATGGCGAGCGAGTGG 780

721 AGCTCTACTTCTTCTGAGGAGCGGCGAGTGGAGTCCGACTGCTATGGCGAGCGAGTGG 780

936 TGGCTGCTGTGGCGCGCTGCTGCAAGGCGGATGAGGCGCGCGACGAGCCCTGAGAGGA 985
DB 1 TGGCTGCTGTGGCGCGCTGCTGCAAGGCGGATGAGGCGCGCGACGAGCCCTGAGAGGA 840

781 TGGCTGCTGTGGCGCGCTGCTGCAAGGCGGATGAGGCGCGCGACGAGCCCTGAGAGGA 840

966 AGTGAACCAAGTCTCTGTAAGAGGCGCGCGTGGAGTGTCTGCGCGCAAGTGGAGCTTACT 1045
DB 1 AGTGAACCAAGTCTCTGTAAGAGGCGCGCGTGGAGTGTCTGCGCGCAAGTGGAGCTTACT 900

841 AGTGAACCAAGTCTCTGTAAGAGGCGCGCGTGGAGTGTCTGCGCGCAAGTGGAGCTTACT 900

1046 TCAACCAAGTCTCTGTAAGAGGCGCGCGTGGAGTGTCTGCGCGCAAGTGGAGCTTACT 1105
DB 1 TCAACCAAGTCTCTGTAAGAGGCGCGCGTGGAGTGTCTGCGCGCAAGTGGAGCTTACT 960

901 TCAACCAAGTCTCTGTAAGAGGCGCGCGTGGAGTGTCTGCGCGCAAGTGGAGCTTACT 960

1106 TTTGGGCTTTTCAAGCAGAGTGGGCTGACATGTACCTGTGCGCATCTGTGATACAGT 1165
DB 1 TTTGGGCTTTTCAAGCAGAGTGGGCTGACATGTACCTGTGCGCATCTGTGATACAGT 1020

961 TTTGGGCTTTTCAAGCAGAGTGGGCTGACATGTACCTGTGCGCATCTGTGATACAGT 1020

1166 TGGGAAGATCCAGGCGGCTTTTGAAGGCGCGCTTATGAAGATACATGAAGAGCGCGAGA 1225
DB 1 TGGGAAGATCCAGGCGGCTTTTGAAGGCGCGCTTATGAAGATACATGAAGAGCGCGAGA 1080

1021 TGGGAAGATCCAGGCGGCTTTTGAAGGCGCGCTTATGAAGATACATGAAGAGCGCGAGA 1080

1226 AGTGGAGCGCTACACTGACCTGTACCAGCGCTCGGCGCTGCTGCTGATTAACACT 1285
DB 1 AGTGGAGCGCTACACTGACCTGTACCAGCGCTCGGCGCTGCTGCTGATTAACACT 1140

1081 AGTGGAGCGCTACACTGACCTGTACCAGCGCTCGGCGCTGCTGCTGATTAACACT 1140

1286 GGCATGCGCGCGCGCGCTACACAGCTCCCTGAGACTACCGGCAACATCTCAACTTGG 1345
DB 1 GGCATGCGCGCGCGCGCTACACAGCTCCCTGAGACTACCGGCAACATCTCAACTTGG 1200

1141 GGCATGCGCGCGCGCGCTACACAGCTCCCTGAGACTACCGGCAACATCTCAACTTGG 1200

1346 TCAAGAACCGCGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1405
DB 1 TCAAGAACCGCGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260

1201 TCAAGAACCGCGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260

1406 TCAAGAACCGCGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1465
DB 1 TCAAGAACCGCGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320

1261 TCAAGAACCGCGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320

1466 CCACCTATACAGTGTCTTCAATTTGGCACAGAGAGCGCTGCTGCTCAAGGCTGTGAGCC 1525
DB 1 CCACCTATACAGTGTCTTCAATTTGGCACAGAGAGCGCTGCTGCTCAAGGCTGTGAGCC 1380

1321 CCACCTATACAGTGTCTTCAATTTGGCACAGAGAGCGCTGCTGCTCAAGGCTGTGAGCC 1380

1526 TGGGCGCGCTGCTGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1585
DB 1 TGGGCGCGCTGCTGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440

1381 TGGGCGCGCTGCTGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440

1586 GCTGCTGCTATCTCAGAGCAAGAGCTGCTTCTGCGGCGCTCCGCTCTCAGCTGCTGTC 1645
DB 1 GCTGCTGCTATCTCAGAGCAAGAGCTGCTTCTGCGGCGCTCCGCTCTCAGCTGCTGTC 1500

1441 GCTGCTGCTATCTCAGAGCAAGAGCTGCTTCTGCGGCGCTCCGCTCTCAGCTGCTGTC 1500

1646 AGCTGCGCGTGGCGAGCTGCAATGAAGTATCGCTGCTGCAAGCTGTCTGCGCGCGG 1705
DB 1 AGCTGCGCGTGGCGAGCTGCAATGAAGTATCGCTGCTGCAAGCTGTCTGCGCGCGG 1560

1501 AGCTGCGCGTGGCGAGCTGCAATGAAGTATCGCTGCTGCAAGCTGTCTGCGCGCGG 1560

1706 ACCCTATTTGGCGCTGAGGCGGCAACACAGCGGCTGTGTGGCGCTGGCTGGCGCACTTGG 1765
DB 1 ACCCTATTTGGCGCTGAGGCGGCAACACAGCGGCTGTGTGGCGCTGGCTGGCGCACTTGG 1620

1561 ACCCTATTTGGCGCTGAGGCGGCAACACAGCGGCTGTGTGGCGCTGGCTGGCGCACTTGG 1620

1766 GATCTTACTGATCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1825
DB 1 GATCTTACTGATCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680

1621 GATCTTACTGATCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680

Query Match Best Local Similarity 98.7%; Score 1617.6; DB 6; Length 3776; Matches 1652; Conservative 0; Mismatches 19; Indels 3; Gaps 2;

LOCUS BD171174 3776 bp DNA Linear PAT 17-JAN-2003

DEFINITION Novel gene and protein encoded thereby.

ACCESSION BD171174

VERSION BD171174.1 GI:27876986

KEYWORDS WO 02052005-A/30.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 3776)

AUTHORS Ohara,O., Nagase,T. and Nakajima,D.

TITLE Patent: WO 02052005-A 30 04-JUL-2002;

JOURNAL KAZUSA DNA RESEARCH INSTITUTE,OSAMU OHARA,TAKAHIRO NAGASE, DAISUKE NAKAJIMA

COMMENT OS Homo sapiens (human)

PN WO 02052005-A/30

PD 04-JUL-2002

PF 20-DEC-2001 WO 2001JP011217

PR 22-DEC-2000 JP 00P 389742

PI OSAMU OHARA,TAKAHIRO NAGASE,DAISUKE NAKAJIMA

PC C12N15/12,C07K14/47

CC Novel gene and protein encoded thereby

FT CDS Location/Qualifiers

source 1..3776

1..3776 Location/Qualifiers

1..3776 /organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT 696 a 1088 c 1191 g 801 t

ORIGIN

1826 GCACTAAGATACAGTACAGGCCACTNCCAAAACATCAGGTGGTGGCGGCACAGACC 1885

1681 CATTATAG--AAAGTCAGGCCACT-CCCAAAAACATCAGGTGGTGGCGGCACAGACC 1737

1886 TGGTGTGGCCCTGCACCTCTCCCTCAAACTGGCCCTGCCGATCCCAACCCCGAGG 1942

1738 TGGTGTGGCCCTGCACCTCTCTCAAACTGGCCCATGGCCCATGGCCCTTTGGGG 1794

854 GGAAGGGCAAGTGTCCCTATGACCCAGCTAAGGGCCATGTGGCTTTGTGATGTGG 913

689 AGCTGTACTGGGCGACACTCAACAACTCTCTGGGACAGGAACCCATTATCTGGTACA 748

914 AGCTGTACTGGGCGACACTCAACAACTCTCTGGGACAGGAACCCATTATCTGGTACA 973

749 TGGGCGCCCGACACTCCATGAAGAGAGTACCTGGCTTTGGCTCAACGAACCTGACT 808

974 TGGGCGCCCGACACTCCATGAAGAGAGTACCTGGCTTTGGCTCAACGAACCTGACT 1033

809 TTGTAGGCTTGCCCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 868

1034 TTGTAGGCTTGCCCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1093

869 TCTACTTTCTTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 928

1094 TCTACTTTCTTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1153

929 CTCGTGTGGCCGCTGTGCAAAAGGCGATATGGGGGGGCGACGGACCTGCAAGAGAGT 988

1154 CTCGTGTGGCCGCTGTGCAAAAGGCGATATGGGGGGGCGACGGACCTGCAAGAGAGT 1213

989 GGACACGCTTCTGAAAGGCGCGCTGGCATGCTGTGCCCGCAACTGGCAGCTTACTTCA 1048

1214 GGACACGCTTCTGAAAGGCGCGCGCTGGCATGCTGTGCCCGCAACTGGCAGCTTACTTCA 1273

1049 ACCAGCTGAGCGCATGTCACACACCTGACAGAGACCTCTGTGCAACACACCTTCTTGG 1108

1274 ACCAGCTGAGCGCATGTCACACACCTGACAGAGACCTCTGTGCAACACACCTTCTTGG 1333.

1109 GGGTTTTTCAAGCAGAGTGGGGGTGATGATGATGATGATGATGATGATGATGATGATG 1168

1334 GGGTTTTTCAAGCAGAGTGGGGGTGATGATGATGATGATGATGATGATGATGATGATG 1393

1169 AAGAGATCCAGCGGCTGTTTGAAGGCGCCCTATGAAGAGTACCATGAGGAAGCCAGAGT 1228

1394 AAGAGATCCAGCGGCTGTTTGAAGGCGCCCTATGAAGAGTACCATGAGGAAGCCAGAGT 1453

1229 GGGACCGGTACAGTACGACCTGATGACCAAGCCCGCGCGCGCGCGCGCGCGCGCGCG 1288

1454 GGGACCGGTACAGTACGACCTGATGACCAAGCCCGCGCGCGCGCGCGCGCGCGCGCG 1513

1289 ATCGCGCCAGCGGCTACACCAAGCTCCCTGAGAGTACCCGACAACTCTCAACTCTGTCA 1348

1514 ATCGCGCCAGCGGCTACACCAAGCTCCCTGAGAGTACCCGACAACTCTCAACTCTGTCA 1573

1349 AGAAGCACCCTGATGAG 1408

1574 AGAAGCACCCTGATGAG 1633

1409 AGAAGGAGACCAACTTCAACCAAGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1468

1634 AGAAGGAGACCAACTTCAACCAAGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1693

1469 CCTATACAGTGTGCTTCAATTTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1528

1694 CCTATACAGTGTGCTTCAATTTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1753

1529 GGGCTGGGCTTACCTGATTTGAGAGCTGACAGCTTTTGGACAGAGAGAGAGAGAGAG 1588

1754 GGGCTGGGCTTACCTGATTTGAGAGCTGACAGCTTTTGGACAGAGAGAGAGAGAGAG 1813

1589 TGGTGTATATCAGAGCAAGAAAGTGTCTTGGCGCGCTTCCCGCTCTCAAGTGTGGAGC 1648

1814 TGGTGTATATCAGAGCAAGAAAGTGTCTTGGCGCGCTTCCCGCTCTCAAGTGTGGAGC 1873

1649 TGCCCGTGGCGAGTGCATTAAGTATGCTCTCTGCAAGAGTGTCTCTGCGCGCGAGCC 1708

1874 TGCCCGTGGCGAGTGCATTAAGTATGCTCTCTGCAAGAGTGTCTCTGCGCGCGAGCC 1933

1709 CCTATGCGCGCTGAGAGCTCAACACACCGCGCTGTGGCGCGTGGGTGGCCACTTTGGAT 1768

Db 1934 CCTATTGCCCTGGACGCTCAACACCCGCTGTGTGGCCGTGGTGGCCACTCTGTGAT 1993
QY 1769 CTTTACGATCAGCATGTGATGATGACCTGAGACCTTCAAGCATTTGCAACCTCCGTGGA 1828
Db 1994 CTCTACTGATCAGCATGTGATGATGACCTGAGACCTTCAAGCATTTGCAACCTCCGTGGA 2053
QY 1839 GTAAGATCAGTCAAGCCCACTNCCCAAAACATACAGGTGTGGCGGACAGACCTGG 1888
Db 2054 GTAAAG-AAAAGTCAAGCCCACT-CCCAAAACATCAGGTGTGGCGGACAGACCTGG 2110
QY 1889 TGCCTCCCTGACCTCTCTCAACTGTTGGCCCTGCGGACCTCAACCCCGGAG 1942
Db 2111 TGTCTCCCTGACCTCTCTCTCAACTGTTGGCCCTGCGGACCTCTTTGGGG 2164

RESULT 10
AB051526 3776 bp mRNA linear PRI 07-FEB-2001
LOCUS Homo sapiens mRNA for KIAA1739 protein, partial cds.
DEFINITION Homo sapiens mRNA for KIAA1739 protein, partial cds.
ACCESSION AB051526
VERSION AB051526.1 GI:12698022
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Nagase,T., Kikuno,R., Hattori,A., Kondo,Y., Okumura,K. and Ohara,O.
TITLE 1 (bases)
XIX. The complete sequences of 100 new cDNA clones from brain which
code for large proteins in vitro
DNA Res. 7 (6), 347-355 (2000)
JOURNAL 21082932
MEDLINE 11214970
PUBMED 2 (bases 1 to 3776)
REFERENCE Ohara,O., Nagase,T. and Kikuno,R.
AUTHORS Direct Submission
TITLE Submitted (22-NOV-2000) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research, 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdna@frc.kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
Fax:81-438-52-3914)
JOURNAL location/Qualifiers
FEATURES
SOURCE
1..3776
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="pJ00464"
/note="vector:pBluescriptII SK plus"
1..3776
/gene="KIAA1739"
/gene="KIAA1739"
/note="Start codon is not identified."
/codon_start=1
/product="KIAA1739 protein"
/protein_id="BAB21830.1"
/db_xref="GI:12698022"
/translation="LCSHLMQPGLSGSCWSEGFPEAGSTHSRLCLLCTLLEAVGSR
KKAABEAKVGMGCPALRPEYPLTLRAALSLMASSGKMLRLVPSFLPAMICLLP
GMRGLRPMGCGGOLRFOKCPILPIRGPHLLVAMKAGSGARLAVRPGSPSA
AMLTPEALATVYVRRFSOTGIDPFLITLTPGLLYVAREBLRFSBALBLDGLAIS
WEAPVEKTECIQKGNKNOTECFNIRLQPYNASHLVYCGTYAROPCTVYNNMLT
LEHGFEDGKGPDPKAGHGLVDELVSATLNFAGTEPILIRNKPISKTE
TLAFELNPHFVGSAYVPSVSGSLVDGDLKLYFFPREHVESDCYAEQVAVAVAVCK
GDMGARLQRMKMTPLKARLASAPMNLVFNOLQAHHTLQDTSNMHTTPEGVQAO
WGDYLSAICEQLEIQRVPGPKYKEHEBKOKDRITDPPSPRPSCTINMIRRH
GTSLELIDNLTNFKNPLMEBOVGFRMSRPLVKGTNTTHLVADRVTGLDQATY
TVLFGTDGMLKAVSLGPNVHLIEQLDQEPMSLVLSQSKRLFPAGRSOLVO
LPVADCKRYRSCADVLARDPYCANVTSRCAVAGSGSLLSQIKHVTSPDSCINL
RGSKRVPRTKIVYAGDLYLPCHLSSNLAHARTGGRDLPAEOGSGFLYDARLO
ALVVAADPRNAGAHGFSBEGARLAAGTYLAVYAVAGPSVTLERALENGLYWLA
VVALGAVCLVLLLVLSLRRLRELELGAATENTLVYPLELPEKPSPPRCPPEP

DEKMDPVGYTYSIDGSLKTIIPGHARCQPGCPPSPGIPGQPLDPSPLRLHGGGRNS.
NANGVRLQLGDEDRGGLHPLPELADLERRLQOROPLPDSNPRESSV*
BASE COUNT 696 a 1088 c 1191 g 801 t
ORIGIN
Query Match 70.8%; Score 1617.6; DB 9; Length 3776;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1652; Conservative 0; Mismatches 19; Indels 3; Gaps 2;
QY 269 CTGGAGAGTGGCCACGAGTATGACGCGGTTCTCCAGACCGGATCCAGACTTCCTTA 328
Db 494 CTGCAAGAGTGGCCACGAGTATGACGCGGTTCTCCAGACCGGATCCAGACTTCCTTA 553
QY 329 CACTGACGCTGACGAGGCCCACTGGGCTTCTGTACCTGGGCGCCGAGAGGCCCTGTTTG 388
Db 554 CACTGACGCTGACGAGGCCCACTGGGCTTCTGTACCTGGGCGCCGAGAGGCCCTGTTTG 613
QY 389 CCTTACGACATGAGAGCCCTGGAGCTGCAAGAGCCATCTCTGCGAGAGGCCCTGGAGAG 448
Db 614 CCTTACGACATGAGAGCCCTGGAGCTGCAAGAGCCATCTCTGCGAGAGGCCCTGGAGAG 673
QY 449 AGAAGACTGATGTATCCAGAAAGGAGACACACAGACGAGTGTCAACTTCATCC 508
Db 674 AGAAGACTGATGTATCCAGAAAGGAGACACACAGACGAGTGTCAACTTCATCC 733
QY 509 GCTTCCCTGAGGCTTCAATATGCTCCACCTGTACGTCTGTGGACCTACGCTTCACAC 568
Db 734 GCTTCCCTGAGGCTTCAATATGCTCCACCTGTACGTCTGTGGACCTACGCTTCACAC 793
QY 569 CCAAGTGGACATGACATGACATGACATGACATGACATGACATGACATGACATGACATG 628
Db 794 CCAAGTGGACATGACATGACATGACATGACATGACATGACATGACATGACATGACATG 853
QY 629 GGAAGGGAAGTGTCTCTATGACCCAGCTAAGGCGCATCTGCGCTTCTGTGATGATG 688
Db 854 GGAAGGGAAGTGTCTCTATGACCCAGCTAAGGCGCATCTGCGCTTCTGTGATGATG 913
QY 689 ACCTGACTGCGGCACACATCAACAACTTCTGGGCAAGGACCATATCTGCGCTTACA 748
Db 914 ACCTGACTGCGGCACACATCAACAACTTCTGGGCAAGGACCATATCTGCGCTTACA 973
QY 749 TGGGGGCCCCACACATCAACAACTTCTGGGCAAGGACCATATCTGCGCTTACA 808
Db 974 TGGGGGCCCCACACATCAACAACTTCTGGGCAAGGACCATATCTGCGCTTACA 1033
QY 809 TTGTAGGCTCTCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 868
Db 1034 TTGTAGGCTCTCTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1093
QY 869 TCTACTTCTTCTTCAAGGAGCGGCGAGTGGAGTCCGACTGCTATGCGGACGAGTGG 928
Db 1094 TCTACTTCTTCTTCAAGGAGCGGCGAGTGGAGTCCGACTGCTATGCGGACGAGTGG 1153
QY 929 CTGCTGTCGCGCTGCTCAAGGGGAGATATGGGGGGGCGGACGAGCCGAGAGGAAT 988
Db 1154 CTGCTGTCGCGCTGCTCAAGGGGAGATATGGGGGGGCGGACGAGCCGAGAGGAAT 1213
QY 989 GGACACAGTCTTCTGAAGGCGGCGTGGCATGCTTGGCCCCGGAAGTGGAGCTCTACTTCA 1048
Db 1214 GGACACAGTCTTCTGAAGGCGGCGTGGCATGCTTGGCCCCGGAAGTGGAGCTCTACTTCA 1273
QY 1049 ACCAGCTGAGGCGATGACACCTGCGAGACACCTCTGGGCAACACACCTTCTTTG 1108
Db 1274 ACCAGCTGAGGCGATGACACCTGCGAGACACCTCTGGGCAACACACCTTCTTTG 1333
QY 1109 GGGTTTTCACACAGTGGGGTGCACATCTACCTGTGGGCGCATCTGTAGTACCACTTTG 1168
Db 1334 GGGTTTTCACACAGTGGGGTGCACATCTACCTGTGGGCGCATCTGTAGTACCACTTTG 1393
QY 1169 AAGAGATCAGAGGGGTGTTGAGGGCCCTATTAAGAGTACATGAGAGCCAGAAAT 1228
Db 1394 AAGAGATCAGAGGGGTGTTGAGGGCCCTATTAAGAGTACATGAGAGAGCCAGAAAT 1453

1719 CTGAGGCTCAACACGACGCGCTGTGTGCGCTGGTGGCCACTTGGATCTTACATCAT 1778
1716 CTGAGGCTCAACACGACGCGCTGTGTGCGCTGGTGGCCACTTGGATCTTACATCAT 1735
1779 CCAGCATGTATGACCTCGGACACTTTCAGCATTTGACACCTCGCTGAGTAAATGACA 1838
1726 CCAGCATGTATGACCTCGGACACTTTCAGCATTTGACACCTCGCTGAGTAAATGACA 1838
1839 GTCAGGCGCCACTTCCCAAAACATCAGCTGTGTGGCGGACAGACCTGTGTGCTGCTG 1898
7314 GTCAGGCGCCACTTCCCAAAACATCAGCTGTGTGGCGGACAGACCTGTGTGCTGCTG 7372
1899 CCACCTCTCTCCCAACTTGGCGCTGCGGACCTCCACACCCGAGC 1942
7373 CCACCTCTCTCCCAACTTGGCGCTGCGGACCTCCACACCCGAGC 7416

RESULT 13
S79463 3503 bp mRNA linear ROD 09-DEC-1995
LOCUS M-Sema F-a factor in neural network development [mice, neonatal
DEFINITION brain, mRNA, 3503 nt].
ACCESSION S79463
VERSION S79463.1 GI:1110598
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3503)
AUTHORS Inagaki,S., Furuyama,T. and Iwahashi,Y.
TITLE Identification of a member of mouse semaphorin family
JOURNAL FEBS Lett. 370 (3), 269-272 (1995)
MEDLINE 95385809
PUBMED 7656991
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI g1bseq 170690] from the original journal article.
This sequence comes from Fig. 1A.
FEATURES
Source location/Qualifiers
1..3503
/organism="Mus sp."
/mol_type="mRNA"
/db_xref="taxon:10095"
1..3503
/gene="M-Sema F"
/note="a factor in neural network development"
100..2604
/gene="M-Sema F"
/note="semaphorin homolog; This sequence comes from Fig.
1A"
1..3503
/codon_start=1
/protein_id="AA35184.1"
/db_xref="GI:1110598"
/translation="MAPHRAVWLLAAGLMLGICAEEMKMMVLPKRYTSGGLVTVVR
FSTGIDFLFLTEHSGILYGARALAFVSEALELQCAISWEPAERKICTOK
GSKNQTECFNFIKLPQNSHLVCGTVAFOPKCTYIMLFTFLDAEFEDGKGP
YDPAKGTGLVGGELYSATLNNFLGTEPLIRKGTGHSKTYLAFMLPEPHVS
AFVDSVSGTGGDDKLYFFSERAVVEYDSEOVNARVAVCGGAGATLTKKT
TFLKARLVGSAIPMKVYFNOLKAVTLKSGVSMHTTEFGYQOARWGMDSACEYOL
EJOIQVREGYKEXSEAOAKMARITDTPYSPRSCINNHRDNGTSSLEPNTLN
FIKHPIMEQVAKRPLVKNKNTNTHVADRVGLDATTVALFISGDMWLK
AVSLGPWIDHVEELQVDPDEVESLVLSQSKVLPASRSQVLQSLADCKRRVVD
CVLARDPICAMANTSRVATTSGRSGSFVQHVANDTSMCKQYGIKVRSLPKNI
TVVSGDLVPLCHLSNLAHMTGSGODLEAEOGSEFLVDTGQALVYMAAQSRHG
PYGYSSEOGTFLAESYLVAVVAGSSVTEARAPLENGLWMLAVVALAGVILL
LVLSIRRLREELKGAASERPLVPLIEPLPAPSPRPGPTDCKMDPVYYS
DSEIKITVGHAKROPGGGPPSPPGITGQPLPSPFRLHLGGRNSNANGIVRLDGGG
DRGSGHPLDELDELRRKLQKROPDPDSNPESSV"

BASE COUNT 726 a 917 c 1057 g 803 t

Query Match 59.2%; Score 1351; DB 10; Length 3503;
Best Local Similarity 84.7%; Pred. No. 1,1e-266;

Matches 1539; Conservative 0; Mismatches 271; Indels 6; Gaps 2;
127 AGCAGCTGGCCCTGAAGCTCAGACCGCGGCTGCGCATGGCCCACTGGGCTGTC 186
61 AGGCTGTGTGACCTGAAGTTTAGACCCCTGGTGTCTGCAATGGCCCACTGGGCTGTC 120
187 TGGCTGTGCGAGCAAGGCTGTGGGCTGGGCAATTGGGCTGAGGTGTGTGAACCTT 246
121 TGGCTGTGCGAGCAAGGCTGTGGGCTGGGCAATTGGGCTGAGGTGTGTGAACCTT 180
247 GTGGCGGTAGAGAGTCTCTTGGGGAGAGTGGCCACAGTATGATAGCGGCTTCTCCAG 306
181 GTGGCGGTAGAGAGTCTCTTGGGGAGAGTGGCCACAGTATGATAGCGGCTTCTCCAG 240
307 ACCGCGATCCAGACTTCTGACACTGACGTGACGAGCCGACCTGGGCTTCTGATGTC 366
241 ACAGGCAATCCAGACTTCTGACACTGACGTGACGAGCCGACCTGGGCTTCTGATGTC 300
367 GGGCGCCGAGAGGCGCTGTGCTTGCCTTGCATGAGAGCGCTGAGCTGCAGAGCCATC 426
301 GGGCGCCGAGAGGCGCTGTGCTTGCCTTGCATGAGAGCGCTGAGCTGCAGAGCCATC 360
427 TCGTGGAGGCGCCCGCTGGAGAGAGAGCTAGTATCCGAGAAAGGAGAACACACAG 486
361 TCGTGGAGGCGCTGGAGAGAGAGCTAGTATCCGAGAAAGGAGAACACACAG 420
487 ACCGAGTCTTCACACTTCCGCTTCTGAGAGCGCTGACATGACCTCCGCTGATGTC 546
421 ACCGAGTCTTCACACTTCCGCTTCTGAGAGCGCTGACATGACCTCCGCTGATGTC 480
547 TGTGGGACCTTACGCTTCCAGCCCAAGTGCACCTACGTCAACATGCTCAGCTGATGTC 606
481 TGTGGGACCTTACGCTTCCAGCCCAAGTGCACCTACGTCAACATGCTCAGCTGATGTC 540
607 GAGCATGAGAGTGTGAAGATGGGAGGCAAGTCTCCCTATGACCCAGCTTAAGGCGCAT 666
541 GAGCATGAGAGTGTGAAGATGGGAGGCAAGTCTCCCTATGACCCAGCTTAAGGCGCAT 600
667 GCTGGCCCTTGTGTGATGATGAGCTGACTGCGCCACACTCAACACTTCTGGGACAG 726
601 ACCGAGTCTTCACACTTCCGCTTCTGAGAGCGCTGACATGACCTCCGCTGATGTC 660
727 GAACCATTAATCTGCGTAACTGATGAGGCGCCCAACATCATGAGACAGATGCTGCGC 786
661 GAGCGGTTATCTGCGTAACTGATGAGGCGCCCAACATCATGAGACAGATGCTGCGC 720
787 TTTGGCTCAACAGAACTCTTGTAGGCTGCTGCTATGACGTGAGAGTGTGGGACG 846
721 TTTGGCTCAACAGAACTCTTGTAGGCTGCTGCTATGACGTGAGAGTGTGGGACG 780
847 TTCACGGGGAGCAGCAGCAAGGCTTCTTCTGAGGAGCGGGCAGTGGAGTCCGAC 906
781 TTCACGGGGAGCAGCAGCAAGGCTTCTTCTGAGGAGCGGGCAGTGGAGTATGAC 840
907 TGTATGCGCAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
841 TGTATGCGCAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
967 GCAGGAGACCTGCGAGAGAACTGAGCAGCAGTCTTCTGAGAGGCGGCTGCAATGCTGCTC 1026
901 GCAGGAGACCTGCGAGAGAACTGAGCAGCAGTCTTCTGAGAGGCGGCTGCAATGCTGCTC 960
1027 CCGAAGTGGAGCTTACTTCAACAGCTGAGGAGGAGCAGACCTTCCAGAGACCTCC 1086
961 CCGAAGTGGAGCTTACTTCAACAGCTGAGGAGGAGCAGACCTTCCAGAGACCTCC 1020
1087 TGGCACAACACACCTTCTGCGGCTTCTGAGAGGCGGAGTGGAGTGGAGTGGAGTGG 1146
1021 TGGCACAACACACCTTCTGCGGCTTCTGAGAGGCGGAGTGGAGTGGAGTGGAGTGG 1080
1147 GCGATCTGTGAGTACCACTTGGAGAGATCCAGAGGCTGTTGAGAGGCGGCTTATGAGAG 1206
1081 GCGATCTGTGAGTACCACTTGGAGAGATCCAGAGGCTGTTGAGAGGCGGCTTATGAGAG 1140

OY	1207	TACACGTGGAGGAGCCAGACAGTGGGAGCCGCTACACCTGACCTGTGATACCCAGCCCTCGGCT	1266	Schaefer, T. E., Brownstein, M. J., Usdin, T. B., Toshiyuki, S., Carninci, P., Prange, C., Kana, S. S., Loquellano, N. A., Peters, G. J., Abramson, R. D., Mullihy, S. J., Bosak, S. A., McEwen, P. J., McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S., Morley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulyk, S. W., Villalón, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shvachenko, Y., Boultard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimmwood, J., Schmutz, J., Myers, R. M., Butlerfield, Y. S., Krzyzanski, M. I., Skalska, U., Smalls, D. E., Scherch, A., Schein, J. E., Jones, S. J., and Harris, M. A.
OY	1208	TGAGAGCCGCCCTGCTGCTGCTGAGAGAGGAGGACCAACTCACCACCTGTGGCGGACCGG	1446	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
DB	1321	TTGGGCGCCGCCCTACTGTTGTAAGAGAAACACTMACTTCAACACAGCTGGTGGCCGACAGG	1380	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
OY	1447	GTTACAGACTTGTATGAGAGCCACCTATACAGTGTCTGTTCAATGGCCACGAGAGAGCGCTGG	1506	Medline
DB	1381	GTCCAGAGGCTGTGATGTGGCCACCTATACAGTGTGTTCAATGGTACAGAGAGATGGCTGG	1440	Pubmed
OY	1507	CTGCTCAAGGCTGTGAGAGCTGGGGCCCTGGGCTTCAACCAGATTGAGAGAGCTCAGCTGTT	1566	Reference
DB	1441	CTGCTGAAGGCTGTGAGAGCTGGGGCCCTGGAGTCCACATGTGTGAGAGAACTGACAGCTGTT	1500	Authors
OY	1567	GACCAGAGCCCATAGAGAGCCGTGCTATCTCAGAGCAAGAAGCTGCTTTGGCCGGC	1626	Title
DB	1501	GACCAGAGAGCCAGTGTGAGAAAGTCTGTCTCTCAGAGCAAGAAAGTCTTTGGCTGGC	1560	Journal
OY	1637	TCCGGCTCTCAGCTGTGTGACAGCTCCCGCGGCCACATGCAATAGATGCTCTGTGCA	1686	Remark
DB	1561	TCCCGCTCTCAGCTGTGTGACAGCTCTCTCTGCGCACATGCAAAAGTACCGTTTCTGTGTA	1620	Comment
OY	1687	GACTGTGTCTCGCGCGCGGAGCCCTATGTGGCGCTGGAGCGTCAACACAGCGGCTGTGTG	1746	
DB	1631	GACTGTGTCTCGCGCGGAGGAGCCCTTACTGTCTGTGCAATGTCAACACAGCGCGTGTGTG	1680	
OY	1747	G---CCGTGGGTGGCCACTTTTGATCTTTACTGTATCCAGATGTGATGACCTCGACACT	1803	
DB	1681	GCCACACACATGTGTGCGCTGCGGGTCTTCTGTGTCMAATGTGGGCAACTGTGGACACT	1740	
OY	1804	TCAGGCAATTGCCAATCTCCCTGGGCAAGTACAGTACAGTACAGGCCCACTNCCCAAAACATC	1863	
DB	1741	TCAAAAGATGTGTACACCACTATGTGCAATTA---AAAAGTCAATCTATTTCCCAAAACATC	1797	
OY	1864	ACGGTGTGTGGGGGACAGACCTGTGTGCTGCTGCCACCTCTCTCCCACTTGGCCCTG	1923	
DB	1798	ACCGTGTGTGTAGACAGACCTGTGTCTACCTGTGCCACCTCTGTCCAAATTTGGCCCAT	1857	
OY	1924	CCCGACTCCAAACCCCG 1939		
DB	1858	GCCCACTGACCTTCG 1873		
RESULT 14				
LOCUS	BC017476	2907 bp	linear	PRI 05-JUN-2003
DEFINITION	Homo sapiens sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4C, mRNA			
ACCESSION	BC017476			
VERSION	BC017476.1			
KEYWORDS	GC			
SOURCE	Homo sapiens (human)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 2907)			
AUTHORS	Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G., Klausner, R. D., Collins, F. S., Wagner, L., Shenmen, C. M., Schuler, G. D., Altschul, S. F., Zeeberg, B., Bueltow, K. H., Schaefer, C. F., Bhat, N. K., Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J., Hsieh, F., Datchenko, L., Marsina, K., Farmer, A. A., Rubin, G. M., Hong, L.,			


```

Db      ||||||| 16 GCTCAACGAACTCTTGTAGGCTCTGCCATATGACTGAGAGTGGGAGCTTCAC 75
OY      852 GGGGAGCAGCAGCAAGTCTACTCTTCTTTCAGGGAGCGGGCAGTGGATCCGACTGCTA 911
Db      76 GGGGAGCAGCAGCAAGTCTACTCTTCTTTCAGGGAGCGGGCAGTGGATCCGACTGCTA 135
OY      912 TCCCGAGCAGTGTGGCTCTGTGGCCCGTGTCTGCAAGGGCGATATGGGGCCGACG 971
Db      136 TCCCGAGCAGTGTGGCTCTGTGGCCCGTGTCTGCAAGGGCGATATGGGGCCGACG 195
OY      972 GACCCGAGCAGAGTGGACAGCTTCTGAAAGGGCGGGCTGGCATGCTGTGCCCGAA 1031
Db      196 GACCCGAGCAGAGTGGACAGCTTCTGAAAGGGCGGGCTGGCATGCTGTGCCCGAA 255
OY      1032 CTGGAGCTCTACTTCAACACAGCTGAGCGATGCAACCTTGAGAGACCTCTGGCA 1091
Db      256 CTGGAGCTCTACTTCAACACAGCTGAGCGATGCAACCTTGAGAGACCTCTGGCA 315
OY      1092 CAACACACCTCTTGGGGTTTTCAGACAGTGGGTGACATGTACCTGTGGCCAT 1151
Db      316 CAACACACCTCTTGGGGTTTTCAGACAGTGGGTGACATGTACCTGTGGCCAT 375
OY      1152 CTGTAGTACAGTGTGGAAAGATCCAGGGGTGTTGAGGGCCCTATAGGAGTACCA 1211
Db      376 CTGTAGTACAGTGTGGAAAGATCCAGGGGTGTTGAGGGCCCTATAGGAGTACCA 435
OY      1212 TGAAGAACCCAGAAAGTGGGACCGGTACACTGACCCTGTACCAGCCCTCGGCTGGCTC 1271
Db      436 TGAAGAACCCAGAAAGTGGGACCGGTACACTGACCCTGTACCAGCCCTCGGCTGGCTC 495
OY      1272 GTGCATTTAACAACTGGCATCGGCCACGGCTACACACAGCTCCCTGGAGCTACCCGACAA 1331
Db      496 GTGCATTTAACAACTGGCATCGGCCACGGCTACACACAGCTCCCTGGAGCTACCCGACAA 555
OY      1332 CATCCCTCACTTCGTCAAGAACCCCGTGTATGGAGAGCAGTGGGGCTTCGGTGGAG 1391
Db      556 CATCCCTCACTTCGTCAAGAACCCCGTGTATGGAGAGCAGTGGGGCTTCGGTGGAG 615
OY      1392 CCGCCCTCTGCTCGTGAAGAAAGGGCACAACTTCAACCCAGCTGTGGCCGACGGGTAC 1451
Db      616 CCGCCCTCTGCTCGTGAAGAAAGGGCACAACTTCAACCCAGCTGTGGCCGACGGGTAC 675
OY      1452 AGCACTTGATGAGCCACCTATACAGTGTGTTCAATTGGCACAGAGACGCTGGCTGCT 1511
Db      676 AGCACTTGATGAGCCACCTATACAGTGTGTTCAATTGGCACAGAGACGCTGGCTGCT 735
OY      1512 CAAGGCTGTGAGCTGGGGCCGTGGGTACAGTGAAGTGAAGAGCTGTGGACCA 1571
Db      736 CAAGGCTGTGAGCTGGGGCCGTGGGTACAGTGAAGTGAAGAGCTGTGGACCA 795
OY      1572 GGAGCCCATGAGAAAGCTGTGCTATCTCAGAGCAAGAGTGTCTTGGCCGGCTCCG 1631
Db      796 GGAGCCCATGAGAAAGCTGTGCTATCTCAGAGCAAGAGTGTCTTGGCCGGCTCCG 855
OY      1632 CTCTCAGCTGTGTGACAGCTGCCGTGGCGAGCTGCATAAAGTATCGCTCTGTGCAGACTG 1691
Db      856 CTCTCAGCTGTGTGACAGCTGCCGTGGCGAGCTGCATAAAGTATCGCTCTGTGCAGACTG 915
OY      1692 TGTCCGCGCGCGGAGCCCTATTGGCGCTGAGCGGCAACACAGCCGCTGTGGCCGT 1751
Db      916 TGTCCGCGCGCGGAGCCCTATTGGCGCTGAGCGGCAACACAGCCGCTGTGGCCGT 975
OY      1752 GGGTGGCCACTTGGATCTTTACTGATCCAGCATGTGATGACCTGCGACCTTCAGAGCAT 1811
Db      976 GGGTGGCCACTTGGATCTTTACTGATCCAGCATGTGATGACCTGCGACCTTCAGAGCAT 1035
OY      1812 TTGCAACCTCCGTGACAGTAAATACAGTCAAGGCCACTTCCCAAAACATCAGCGTGT 1871
Db      1936 CTGCAACCTCCGTGACAGTAAAGTCAAGGCCACTTCCCAAAACATCAGCGTGT 1092
OY      1872 GGGGGGACAGACCTGTGTGCTCCCTGGCACCTTCCTCAACTTGGCCCTGCCGACTTC 1931

```

```

Db      1093 GGGGGGACAGACCTGTGTGCTGCCCTGCCACCTTCCTCAACTTGGCCCATGCCCGCTG 1152
OY      1932 CAACCCCGAGG 1942
Db      1153 GACCTTGGGG 1163

```

Search completed: August 19, 2003, 14:58:53
 Job time : 5567 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2003, 13:45:35 ; Search time 43 Seconds

(without alignments)
1332.943 Million cell updates/sec

Title: US-10-002-050-14

Perfect score: 3219

Sequence: 1 MAPHNAVMLAARLMLGLIG.....PCHLSSNLAIPDSNPRESSV 596

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen param

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : PIR76:*

1: pirl1:*
2: pirl2:*
3: pirl3:*
4: pirl4:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2720	84.5	834	2	M-sema F protein p
2	1149	35.7	782	2	semaphorin C - mou
3	992.5	30.8	760	2	semaphorin B - mou
4	934.5	29.0	772	2	semaphorin - chicke
5	932.5	29.0	771	2	semaphorin III pre
6	931.5	28.9	772	2	semaphorin D - mou
7	892	27.7	749	2	semaphorin V - hum
8	872.5	27.1	748	2	semaphorin A - mou
9	858	26.7	666	2	semaphorin III - m
10	842	26.2	753	2	semaphorin III fem
11	774	24.0	751	2	semaphorin E - mou
12	739	23.0	712	2	hypothetical prote
13	729.5	22.7	1074	2	semaphorin IV precu
14	708.5	22.0	730	2	semaphorin F precu
15	685	21.3	724	2	semaphorin II prec
16	658	20.4	656	2	semaphorin I - fru
17	597	18.5	711	2	semaphorin I - fru
18	406.5	12.6	653	2	semaphorin I - fru
19	368	11.4	676	2	semaphorin homolog
20	236.5	7.3	1945	2	hypothetical prote
21	213	6.6	1894	2	plexin A - fruit f
22	207	6.4	1905	2	plexin 1 precursor
23	198	6.2	1872	2	plexin 3 precursor
24	196.5	6.1	1884	2	plexin 2 precursor
25	186	5.8	403	2	A39R protein - vac
26	183	5.7	441	2	hypothetical prote
27	166.5	5.2	1568	2	semaphorin recept
28	138.5	4.3	2051	2	plexin B - fruit f
29	118.5	3.7	620	2	hypothetical prote

30	116.5	3.6	1375	1	JC5148	hepatocyte growth
31	116	3.6	1425	2	T30811	hepatocyte growth
32	115	3.6	295	2	J01775	saliv protein - v
33	114.5	3.6	1806	2	T23298	hypothetical prote
34	108.5	3.4	1047	2	B71402	hypothetical prote
35	107	3.3	1379	1	S01254	hepatocyte growth
36	102	3.2	577	2	A37779	histocompatibility
37	99.5	3.1	860	2	F71000	hypothetical prote
38	98.5	3.1	1154	2	S39536	paraspinal crystal
39	97.5	3.0	523	2	T04825	hypothetical prote
40	96	3.0	552	2	AD1824	periplasmic oligop
41	96	3.0	1503	2	T18266	cyclinulo-oligosa
42	95	3.0	1926	2	S01169	beta-glycosidase c
43	94	2.9	1428	2	C85079	hypothetical prote
44	94	2.9	1607	2	T03022	MAP kinase kinase
45	93.5	2.9	684	1	HJECDA	helicase (EC 3.6.1

ALIGNMENTS

RESULT 1

S66498 M-sema F protein precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 24-Nov-1999

C:Accession: S66498

R:Inagaki, S.; Furuyama, T.; Iwahashi, Y.

FEBS Lett. 370, 269-272, 1995

A:Title: Identification of a member of mouse semaphorin family.

A:Reference number: S66498; MID:95385809; PMID:7656991

A:Accession: S66498

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-834 <IN>

A:Cross-References: EMBL:S79463; NID:g1110598; PIDN:AAB35184.1; PID:g1110599

C:Superfamily: semaphorin

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-834/Product: M-sema F protein #status predicted <MAT>

Query Match	Score	Length	DB 2;	Length	DB 34;
Best Local Similarity	84.5%	834;			
Matches 501; Conservative	41;	Mismatches 47;	Indels 14;	Gaps 4;	
OY	1	MAPHNAVMLAARLMLGLIGAEVMMNLVPRKTVSSGELATVVRFSOTGIODPLTFTTE	60		
DB	1	MAPHNAVMLAARLMLGLIGAEVMMNLVPRKTVSSGELATVVRFSOTGIODPLTFTTE	60		
OY	61	PTGLLYVGAREALFAFSMEALDELQGAISWEAFVEKTECIQKNNQTECFNFIREFQPY	120		
DB	61	HSGLLYVGAREALFAFSMEALDELQGAISWEAFVEKTECIQKNNQTECFNFIREFQPY	120		
OY	121	NASHLYVCGTFAFORKCTVNNMLTFTLEGEREDGKPKPYPAAGHALLVDGEIYSAT	180		
DB	121	NASHLYVCGTFAFORKCTVNNMLTFTLEGEREDGKPKPYPAAGHALLVDGEIYSAT	180		
OY	121	NSSHLYVCGTFAFORKCTVNNMLTFTLEGEREDGKPKPYPAAGHALLVDGEIYSAT	180		
DB	121	NSSHLYVCGTFAFORKCTVNNMLTFTLEGEREDGKPKPYPAAGHALLVDGEIYSAT	180		
OY	181	LNNFLGTEPIILRNKQPHMSKTEYLAFNLNPHFVGSAYVESGFTGDDDKYFFER	240		
DB	181	LNNFLGTEPIILRNKQPHMSKTEYLAFNLNPHFVGSAYVESGFTGDDDKYFFER	240		
OY	181	LNNFLGTEPIILRNKQPHMSKTEYLAFNLNPHFVGSAYVESGFTGDDDKYFFER	240		
DB	181	LNNFLGTEPIILRNKQPHMSKTEYLAFNLNPHFVGSAYVESGFTGDDDKYFFER	240		
OY	241	ERAVESDCYAEQVAVARVAVRCGDMGARTLQKRTTFLKARLACAPAMQOLYFNOLQAM	300		
DB	241	ERAVESDCYAEQVAVARVAVRCGDMGARTLQKRTTFLKARLACAPAMQOLYFNOLQAM	300		
OY	241	ERAVESDCYAEQVAVARVAVRCGDMGARTLQKRTTFLKARLACAPAMQOLYFNOLQAM	300		
DB	241	ERAVESDCYAEQVAVARVAVRCGDMGARTLQKRTTFLKARLACAPAMQOLYFNOLQAM	300		
OY	301	HLQDTSWNTTFFGVFOQKQNDATLSAICEYOLEIQAVFGPPTKEITEEAKQNDRYTD	360		
DB	301	HLQDTSWNTTFFGVFOQKQNDATLSAICEYOLEIQAVFGPPTKEITEEAKQNDRYTD	360		
OY	301	HLQDTSWNTTFFGVFOQKQNDATLSAICEYOLEIQAVFGPPTKEITEEAKQNDRYTD	360		
DB	301	HLQDTSWNTTFFGVFOQKQNDATLSAICEYOLEIQAVFGPPTKEITEEAKQNDRYTD	360		
OY	361	PVPSRPSCINNNRRHGRGTSLSLEPDLILNFVKKHPLMEBOVGRBMSRPLVKKGTNF	420		
DB	361	PVPSRPSCINNNRRHGRGTSLSLEPDLILNFVKKHPLMEBOVGRBMSRPLVKKGTNF	420		
OY	361	PVPSRPSCINNNRRHGRGTSLSLEPDLILNFVKKHPLMEBOVGRBMSRPLVKKGTNF	420		
DB	361	PVPSRPSCINNNRRHGRGTSLSLEPDLILNFVKKHPLMEBOVGRBMSRPLVKKGTNF	420		
OY	421	THLVADRYTGLDGAATYVLFITGTGQWLLKAVSLCPWVHLIBELQLFDOEPKRSIVLSOS	480		
DB	421	THLVADRYTGLDGAATYVLFITGTGQWLLKAVSLCPWVHLIBELQLFDOEPKRSIVLSOS	480		

[illegible]

RESULT 2

semaphorin C mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 03-Nov-2000
C:Accession: I48746
R:Puschel, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A:Reference number: I48744; MUID:95267431; PMID:7748561
A:Accession: I48746
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-782 <RES>
A:Cross-references: EMBL:X85992; NID:g854327; PIDN:CNA5984.1; PID:g854328
C:Genetics:
A:Gene: semc
C:Superfamily: semaphorin

Query Match	35.7%	Score 1149;	DB 2;	Length 782;
Best Local Similarity	43.6%	Pred. No. 1.1e-86;		
Matches 247; Conservative	89;	Mismatches 201;	Indels 30;	Gaps 11;

```

0Y 41 VVRBR5QSGIDDFLVLTLTEPGLLYGARELLFPM5MAL5ELOGA-----ISM6PVR5K 96
0Y 4 LIRFEANISNTYTLALLS0SGKTLTYGARELLFPLNSLSTLPG6EY0ELMSADNRK 63
0Y 97 TEOIGKKN0T ECE5NFI7FLOPYNASHLYVGYTAFOPKCTYVNMJ7TELEHGE----- 151
0Y 64 OOSGSKGDPKRDCCNYIKILPLPLNSSHLLTGTGA5F5LCAV1HIA5FTLLAQDEAGNYI 123
0Y 152 FEDGKGRCPYAPAKHAGLLVDGELYSANTLNNEFTERTPLIIBNMGRH5MKREYLAFTN 211
0Y 124 LEDGKHC5PDPNFKSTALVYDGLYTGVS5FQNDPIS5QSRRPT-KTES5LNMWQ 182
0Y 212 EPHFVGA5VE5VSG5FTGDDDKVYFFERERAVESDCA5EOVAVARVARKC0DMGARTL 271
0Y 183 DPAFVASAT5E5SL5PGDDDKIYFF5SETQ5EPEF5ENTIV5RAVARKC0DEGERVL 242
0Y 272 ORKRTTFLKACLAGAPRMQOLYFNOLAMHTQDPR-SMHTTFFG5VYQAO5R-5DMYLS 327
0Y 243 QORNT5FLKALCL5SRPDG5RPENLVADVFTLNPMDKRTKL5IGVFT5OHNRGCTEG5 302
0Y 328 AICEV0LEI0RVE5GPYKEVHEEAKMDRYDPP5SPR5G5CINNMRH5RGYT5LELP 387
0Y 303 AICVFTMDVQKAFGLDKVNR5RETQ5WYTEHQVPTPR5GACITNSA5ERKINS5LOLP 362
0Y 388 DNIINFVKH5HLM5EQ5VCPKMSRPLLYVKKGTFT5LADRYG5LDG5ATVYTF5IG5GDM 447
0Y 363 DRYV5FLKDH5FLMD5QVR---5RLL5LOPRARYQ5VA5H5RVGLH-STYDVL5FLG5GDR 418
0Y 448 LKAV5L5PWNH5LEI5LOL5DFO-EP5R5VL5YS05SK5LLF5AG5RSOL5VOL5PVD5CIKYR5C 506
0Y 419 LHKAVT5LSRVH5IIE5LOI5PQ5G5QPV5QW5LL5DSHG5GLLY5ASH5G5VQ5VP5AN5CSLYPTC 478
0Y 507 ADCVL5ARDPY5KANS5NT5RCV5AV5G5H5F5G5SL5LI5QHT5V5DTS5G-----I5NL5G5K---I5 558

```

```

Db      479 GDCLLARPDPYCAWTSAGCISL-----YQDLASRPMTQIEGASVRELCNSSYKARFLV 535
Qy      559 SGPXPKRNTIVAGCTDLVLPCHLSNLA 585
       | : : : | | | | |
Db      536 PGRPKQVQIOPNTVNTTACPLPLSNLA 562

```

RESULT 3

semaphorin B mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Nov-1999
C:Accession: I48745
R:Puschel, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creat
A:Reference number: I48744; MUID:95267431; PMID:7748561
A:Accession: I48745
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-760 <RES>
A:Cross-references: EMBL:X85991; NID:g854325; PIDN:CAA59983.1; PID:g854326
C:Genetics:
A:Gene: semb
C:Superfamily: semaphorin

Query Match	30.8%	Score 992.5;	DB 2;	Length 760;
Best Local Similarity	39.9%	Pred. No. 9.6e-74;		
Matches 228;	Conservative 91;	Mismatches 210;	Indels 43;	Gaps 15;

[illegible]

A49069

collapse - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 24-Sep-1999
 C:Accession: A49069
 R:Luo, Y.; Raper, J.A.
 Cell 75, 217-227, 1993
 A:Title: Collapse: a protein in brain that induces the collapse and paralysis of neuron
 A:Reference number: A49069; MUID:94006554; PMID:8402908
 A:Accession: A49069
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-772 <LUD>
 A:Cross-references: GB:002528; NID:9410078; PIDN:AAC59638.1; PID:9410079
 C:Superfamily: semaphorin

Query Match 29.0%; Score 934.5; DB 2; Length 772;
 Best Local Similarity 38.2%; Pred. No. 6.2e-69;
 Matches 207; Conservative 93; Mismatches 20; Indels 37; Gaps 14;

QY 8 WLALRLMGIG-----IGAEVWNLVPRKTVSSGELA--TVARRFSOTGIODELTLTL 58
 DB 3 WLRGIALSLGLVLLAGRNCOHKNVPRKLSYKEMESNNVIFNGLANSSVHTFL 62

QY 59 TEPGLLYGAREALFAFSMEALDELQALISWAPVEKTECIQKKNNOTECFNTRFLQ 118
 DB 63 DEERSRLVYGAKDIFSFNLNKEYOKIWPVSHSRDECKMAGKDLRECCANIKYK 122

QY 119 PYNASHLYVCGTYAPQPKCTVYNNLT-----FTLEHGEFEDGKCPYDPAKGAGLLV 172
 DB 123 TYNGTHLYACGTGAFHPICTYIEGSHPEDNIFKEDSHFENGKRSKSPYDKLTLASLLY 182

QY 173 DGLYSATLNNFLTEPIILRNMGPHHSKTE-YIAFWLNEPHVGSAYVPSVGSFTGD 231
 DB 183 DGLYSATLADFMGRDFAIFRTIGHNHPRIREQHDSRWLNDPRFISAHILPESDNP--E 239

QY 232 DDKYFFPREAREVSDCAEOVAVARVCKGDMGATLQKRTTFKARLACAP--- 288
 DB 240 DDKYFFPREAREVSDCAEOVAVARVCKGDMGATLQKRTTFKARLACAP--- 288

QY 289 NMOLYFNOLAMHTLQDTSMHTTFEGVFOAQMGDMYLSAICEYOLEEIQRYFEGPYKEX 348
 DB 300 GIDHFEDELQDFVFLNFKNDPRNPVYGVFTSSNIFKSAVCMYMSDVRRTFLGPRYAH 359

QY 349 HEBAOKDRYTDPPSPRPGSCINNMHRHGYTSSLELDPNLIINLVKHPIMEEOVGRPW 408
 DB 360 DGPYQWVPYOGVRVPRGTCPSKTF--GGFDSYKIDLPDEYITFARSHPRMYNVPFIN 417

QY 409 SRPLLVRKGTN--FTHLVADRVTGLDCAATYVLFITGDSGLTKAVSL--GPVH----L 460
 DB 418 SRPLMKTIDVYQFQIYVDRDAEDG-QYDVMFIDGIVLAKVYSIPKETWHELEVYL 476

QY 461 IEEIQLFDOEP--MRSVLVLSQSKLLFAGSRQSLVOLPVAICIKY-RSCADCVLARDPYC 517
 DB 477 LEEMTVF-REPTVISAMKISIRKQOLYIGSATVSQLPLHRDIDYGRKACACBCLARDPYC 535

QY 518 AW 519
 DB 536 AW 537

RESULT 5
 D49423
 semaphorin III precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-Sep-1999
 C:Accession: D49423
 R:Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.
 Cell 75, 1389-1399, 1993
 A:Title: The Semaphorin genes encode a family of transmembrane and secreted growth cone
 A:Reference number: A49423; MUID:94094332; PMID:8269517
 A:Accession: D49423
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA

A:Residues: 1-771 <KOI>
 A:Cross-references: GB:L6081; NID:9799328; PIDN:AAA65938.1; PID:9436560
 C:Genetics:
 A:Gene: SEMA1
 C:Cross-references: GB:283448
 C:Superfamily: semaphorin

Query Match 29.0%; Score 932.5; DB 2; Length 771;
 Best Local Similarity 38.4%; Pred. No. 9e-69;
 Matches 208; Conservative 94; Mismatches 20; Indels 37; Gaps 15;

QY 8 WL--LAARLWIGIGAEVW---NLVPRKTVSSGELATVARRFSOTGI--QDFLTLTL 58
 DB 3 WLRIVCLFWGVLLTARANYONGKNVPRKLSYKEMESNNVIFNGLANSSVHTFL 62

QY 59 TEPGLLYGAREALFAFSMEALDELQALISWAPVEKTECIQKKNNOTECFNTRFLQ 118
 DB 63 DEERSRLVYGAKDIFSFNLNKEYOKIWPVSHSRDECKMAGKDLRECCANIKYK 122

QY 119 PYNASHLYVCGTYAPQPKCTVYNNLT-----FTLEHGEFEDGKCPYDPAKGAGLLV 172
 DB 123 AYNGTHLYACGTGAFHPICTYIEGSHPEDNIFKEDSHFENGKRSKSPYDKLTLASLLI 182

QY 173 DGLYSATLNNFLTEPIILRNMGPHHSKTE-YIAFWLNEPHVGSAYVPSVGSFTGD 231
 DB 183 DGLYSATLADFMGRDFAIFRTIGHNHPRIREQHDSRWLNDPRFISAHILPESDNP--E 239

QY 232 DDKYFFPREAREVSDCAEOVAVARVCKGDMGATLQKRTTFKARLACAP--- 288
 DB 240 DDKYFFPREAREVSDCAEOVAVARVCKGDMGATLQKRTTFKARLACAP--- 288

QY 289 NMOLYFNOLAMHTLQDTSMHTTFEGVFOAQMGDMYLSAICEYOLEEIQRYFEGPYKEX 348
 DB 300 GIDHFEDELQDFVFLNFKNDPRNPVYGVFTSSNIFKSAVCMYMSDVRRTFLGPRYAH 359

QY 349 HEBAOKDRYTDPPSPRPGSCINNMHRHGYTSSLELDPNLIINLVKHPIMEEOVGRPW 408
 DB 360 DGPYQWVPYOGVRVPRGTCPSKTF--GGFDSYKIDLPDEYITFARSHPRMYNVPFIN 417

QY 409 SRPLLVRKGTN--FTHLVADRVTGLDCAATYVLFITGDSGLTKAVSL--GPVH----L 460
 DB 418 NRPYIKTIDVYQFQIYVDRDAEDG-QYDVMFIDGIVLAKVYSIPKETWHELEVYL 476

QY 461 IEEIQLFDOEP--MRSVLVLSQSKLLFAGSRQSLVOLPVAICIKY-RSCADCVLARDPYC 517
 DB 477 LEEMTVF-REPTVISAMKISIRKQOLYIGSATVSQLPLHRDIDYGRKACACBCLARDPYC 535

QY 518 AW 519
 DB 536 AW 537

RESULT 6
 I48747
 semaphorin D - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
 C:Accession: I48747
 R:Puschel, A.W.; Adams, R.H.; Betz, H.
 Neuron 14, 941-948, 1995
 A:Title: Murine semaphorin D/collapse is a member of a diverse gene family and creat
 A:Reference number: I48747; MUID:95267431; PMID:7748561
 A:Accession: I48747
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-772 <RES>
 A:Cross-references: EMBL:X85993; NID:9854329; PIDN:CAAS985.1; PID:9854330
 C:Genetics:
 A:Gene: semd
 C:Superfamily: semaphorin

Query Match 28.9%; Score 931.5; DB 2; Length 772;
 Best Local Similarity 38.6%; Pred. No. 1.1e-68;

C:Genetics:
A:Gene: semaphorin
C:Superfamily: semaphorin

Query Match 24.0%; Score 774; DB 2; Length 751;

Best Local Similarity 33.5%; Pred. No. 1,1e-55;
Matches 155; Conservative 106; Mismatches 187; Indels 34; Gaps 13;

```

OY 52 DEFTLTLEPTGLLYVGAREALFAFMEALFLOG-AISWAPVEKTECIQKKNQTEC 110
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 53 DYRIILMEDEDRIVYGSKDHLISLNNISQEPISLVFNPASTIKVEECMAKDPDTHGC 112
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 111 ENFIRLOPYNASHLYVCGTYAFQPKCTYNN-----LFTLEHEFEDGKCKCYDPA 164
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 113 GNEFVRIQFNRIHLVYCGSAGASPCTYLNRCRSEDOVFMD-SKCEGKCRCSFNPN 171
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 165 KCHAGLLVGCGLYSATLNNFLGEPITILRNMGPHHSNKT-EYLAFLNBEHFVGSAYVPE 223
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 172 VNTVSMINEBELFSGKITDPKGTDAIFSLTRMQLRFDQHSKSLSEPMFADAHYIPD 231
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 224 SVGSFTGDDDKYFFPRERAVESDCYAEQVAVARVAVCKDMGARTLQKWTFLKARL 283
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 232 GEDP---NDAKYVFEKERLTDNNRSTKQIHSWIARICPDGTGQSLVKNKTTFLKARL 288
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 284 ACSAPRW---OLYFNOLQAMHTLQDTSNNTFFPGVFOQMGDMYLSAICEYOLEIORY 340
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 289 VCSVTDEDEPETHFDELEDFLEETDNPRTTLVGLFTSSSVFKSACVYHLSIDQYV 348
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 341 FEGPYKVEHEAOKMDRYTDPVSPRPGSCINNMHRHGYTSL---ELPDNILNPFVK 396
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 349 FNGPFAHKGSPNQLGSLYSGRIYPRGCPG-----AFTPMRTTKDPEDDVVFIRN 403
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 397 HPLMEQVGRMSRPLLVKKGIN--FTHLVADRYTGLDCAATYVLFITGEGDMLKAV-- 452
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 404 HPLMYSISPIHRRPLIVRGTDYKTKYKIAVDVNMADG-RYHVLFGTGRGTVOQRYVVL 462
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 453 ---SLCPWHLIEQLLF-DOEPMSLAVYSQSKLLFAGRSQDLNOLPVADCIKY--RSCA 507
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 463 PNTSSASGELLELEEFKNHPTTMTETSSKKOOLVSSNEGVSQSLRCHYGTACA 522
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 508 DCVLARDPYCAM 519
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 523 DCCLARDPYCAM 534
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

RESULT 12
T27165
hypothetical protein Y545B.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27165
R:Kennard, N.
submitted to the EMBL Data Library, October 1998
A:Reference number: 220321
A:Accession: T27165
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-712 <M12>
A:Cross-references: EMBL:AL03653; PIDN:CAA21714.1; GSPDB:GND00019; CESP:Y545B.1
A:Experimental source: clone Y545B
C:Genetics:
A:Gene: CESP:Y545B.1
A:Map position: 1
A:Introns: 36/1; 70/2; 85/3; 194/2; 221/1; 253/3; 301/1; 352/3; 522/2; 560/1; 599/3; 655

Query Match 23.0%; Score 739; DB 2; Length 712;
Best Local Similarity 34.0%; Pred. No. 8.2e-53;
Matches 133; Conservative 96; Mismatches 214; Indels 64; Gaps 20;

OY 26 NLVPRVSSGELATVVRFSQTCI---ODEFTLTLEPTGLLYVGAREALFAFMEAL 81
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 26 NLVPRQIINS---VGIGDRFGGIGTSSDSDESHKLLAADDSDLL-VGARNAVYINLSLTL 81
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

OY 82 ELQGAISWEAPVEKTECIQKKNQTECNEFTRELOPYNASHLYVCGTYAFQPKCTYVN 141
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 82 SVNHKIDMKPRAPRIECCIMKGR-SKPDCONYIRVLARKSAGSLVCGTAHSPKCREYT 140
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 142 MLPTLEHGEFEDGKCKCPYDPAKAGLLVDG--ELYSATLNNFLGTEETLILR-NMGPH 198
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 141 VTEFGIKNTROFPGQGLSPYDPKHNSSALVPTGTNOLVATVDFGNDLLIRKTTDET 200
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 199 HSNKT-----EYLAFLNBEHFVGS-AVPSVGSFTGDDDKYFFPRERAVES-DCY 249
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 201 PSSKSAANIRTOGYDARVLANAPVATFAI-----KHEHYWFREINSEADLN 249
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 250 AE-QVAVARVAVCKDMGARTLQKWTFLKARLACAPNMO--LYFNOLQAMHTLQDT 306
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 250 EEPQIYARVAVRCKNKGARPANERTSYLAKRLNCLSPGSPFEYFLKAVSDPIDA 309
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 307 SWNHTFFPGVFOQMGDMYLSAICEYOLEIOWFE-GPKVEHEAOKMDRYT-DPVPS 364
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 310 GNNHVVYTVFSTPPSDVRRSAVCKFSMKIRIEFDGTYRKHONNOSMMAFRNREVPK 369
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 365 PRPGSCINNMHRHGYTSLLEPDNILNPFVKHPLMEQVGRMSRPLLVKKG--TYFTH 422
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 370 PRPGSC-----SPDSRKLPENTVVSFLHHPHLHRL-PSVAAPLVLVEGADRADLYQ 419
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 423 L-VADRYTGLDCAATYVLFITGEGDMLKAVSLGPRVHLIEQLFQDEPRMSLVLSQSK 461
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 420 ITVLPRAVAGNHYDLFTIGTSDGKVLKVEYDAGNATVQSAVTF-QRGVPYVNLITTK 478
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 482 KLFAGRSQDLNOLPVADCIKYRSCADCVLARDPYCAMSVNTSCVAVGHPSLTQYH 541
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 479 ESVVIYSADIASLAPVINCAGQTSCKSCYVQLQDPHCAMDSIARCVHSGSWTDQFIQM 538
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 542 MTSDTSCICNLKSGSKIOSGPRKNITV 568
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 539 VFG-----QSEQCPEGIIV 552
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

RESULT 13
JC5928
semaphorin F precursor - human
C:Species: Homo sapiens (man)
C>Date: 10-Apr-1998 #sequence_revision 08-May-1998 #text_change 17-Nov-2000
C:Accession: JC5928
R:Simmons, A.D.; Puschel, A.W.; McPherson, J.D.; Overhauser, J.; Lovett, M.
Biochem. Biophys. Res. Commun. 242, 685-691, 1998
A:Title: Molecular cloning and mapping of human semaphorin F from the Crl-du-chat can
A:Reference number: JC5928; MUID:8135554; PMID:9464278
A:Accession: JC5928
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1074 <SIM>
A:Cross-references: GB:052840; NID:g2772583; PIDN:MAC09473.1; PID:g2772584
A:Experimental source: brain
C:Comment: This protein disrupts normal brain development and leads to some of the fe
C:Genetics:
A:Gene: semaf
C:Superfamily: human semaphorin F; thrombospondin type 1 repeat homology
F:1-20/Domain: signal sequence #status predicted <SIG>
F:50-533/Domain: semaphorin #status predicted <SEN>
F:840-896/Domain: thrombospondin type 1 repeat homology <THR>
F:971-993/Domain: transmembrane #status predicted <TM>

Query Match 22.7%; Score 729.5; DB 2; Length 1074;
Best Local Similarity 32.7%; Pred. No. 9e-52;
Matches 188; Conservative 81; Mismatches 219; Indels 87; Gaps 21;

OY 8 WLAARLWGIGIGAEVWNL-----VPRKTVSSGELATVVRFSQTCIODEFL 54
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 9 WLFSS-----LGL-----WRLAHPEAOCTTQCOPTHFVVISYKEIGPWLRFPRAKNADFS 59
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

OY 55 TLFTLEPTGLLYVGAREALFAFMEALFLOGAISWEAPVEKTKKTECIQKKNQTECNEFT 114
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 60 OLTFDEQKELVAGARNYFLRIQLEDSLQIOWAVECDEATKACYSKGSKE-ECONYI 118
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY 115 RFLQPYASHLIYCGCTAFOPKCTYVNMILFTLEHGEFEDGKCPYDPANGHAGLL-VD 173
 Db 119 RVLV-VGGDRLEFGTNAFTPEVCNRSLSNLAEIHDL-SGMARCPYSPQHNSTALLNAG 176
 QY 174 GELYSATLNNNLGEPILIRNMGPHHSMT-EYLAFLNNEPHFGSATVPESVSGFTGDD 232
 Db 177 GELTAAATAMDPGGDPALYRSLGILPLRLTAQYNSKMLNEPNVSS---YDIGFT--- 229
 QY 233 DKYFFERERAVEIDCYAEQVAVARVCCKMGCAFTLORNTTFLKARLACASAP-NMQ 291
 Db 230 ---YFFERENAVEIDC-GKTYFSAARVCCKNDIGRFLLEDTTTFMARLNCSPGCEVP 285
 QY 292 LYFNQIQAMHTLQDTSMHNTTFPGVFOAQMGDMYLSAICEYOLEIQRVFBGPKEVHEE 351
 Db 286 FYYNELQSTFELPELD---LYIGFLTNNVNSIASAVCFNLSLAQAFSGPFGYQNS 341
 QY 352 AQKMDRTDPPSPRPGSCINNMHRHGTSS---LELPDNLNFKVKKHPLMEQVGRW 408
 Db 342 RSAN---LPYPNPNP-----HFQCTVQGLVYNTLERNLDAQKFLVHEVVOQVPT 390
 QY 409 SRPLVKKGTFTFLVADRYTGLDQATYVLFEGNGDMILKA-VSLGPMVH---LIEELQ 465
 Db 391 TVPSPMEDNSRFSHVADYVQGRALVH-IYTLATDYGTIKKRVPLNQTSSCLLEIE 449
 QY 466 LPDQ---EPMSIVLSQSKLLFAGSRQVQLVADCIKXKSCADYALADPYCAMSUN 522
 Db 450 LFERREPRIRSLQILHSQSVLFGRLREHYVKIPLKRCQFYRTSTCIGADDPYCGMVOY 509
 QY 523 TSRC-----VAVGHRG 534
 Db 510 MKCTSLSELSMTQWQEOSISACPTRLNLTVDGHEG 544

RESULT 14

JH0798
 fasciclin IV precursor - American bird grasshopper
 C:Species: Schistocerca americana (American bird grasshopper)
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
 C:Accession: JH0798
 R:Kolodkin, A.; Mathies, D.J.; O'Connor, T.P.; Patel, N.H.; Admon, A.; Bentley, D.; Good
 Neuron 9, 831-845, 1992
 A:Title: Fasciclin IV: Sequence, expression, and function during growth cone guidance in
 A:Reference number: JH0798; MUID:93040225; PMID:1418998
 A:Accession: JH0798
 A:Molecule type: mRNA
 A:Residues: 1-730 <KOL>
 A:Cross-references: GB:L00709; MID:g160844; PID:g160845
 A:Experimental source: embryo
 C:Comment: This protein plays a role in growth cone guidance in the developing central n
 C:Keywords: glycoprotein; transmembrane protein
 F:1-22/Domain: signal sequence #status predicted <SIG>
 F:23-730/Product: fasciclin IV #status predicted <KAT>
 F:23-621/Domain: extracellular #status predicted <EXT>
 F:628-652/Domain: transmembrane #status predicted <TM>
 F:653-730/Domain: intracellular #status predicted <INT>
 F:44,71,163,267,360,539/Binding site: carboxydrate (Asn) (covalent) #status predicted

Query Match 22.0%; Score 708.5; DB 2; Length 730;
 Best Local Similarity 31.3%; Pred. No. 2.8e-50;
 Matches 190; Conservative 107; Mismatches 245; Indels 65; Gaps 21;
 QY 6 AVMLLAARLNGIGIGAEVWNLV-PRKTVSSGELATVVRFSQIGQDFLTLEPTGL 64
 Db 4 ALVAVALLW-VALLHAAMVNDVSPKMYVQGE-ERNQRFIGNESHKHFFLEDNHSL 61
 QY 65 LYVCARELFAFSSEALE--LQGAISWEAPYEKTECIQKGNKNOTECFNIRFLQPYNA 122
 Db 62 L-VGARNVIVYINISRLDLETFEQRILEMSSGAHRELCTLKGR-SEDDCONYIRVLAKIDD 119
 QY 123 SHLVVCGTYAFQPKCTYVNMILFTLEHGEF-----EDGKGCPRDPANGHAGLLVDGLY 177
 Db 120 DRVLICGNMAYKPLCRH-----VALKQGDYVERKETBERGLCPDPDHNSTAIYSEGLY 174

QY 178 SATLNFLOTETPIILIRNMGPHHSMTLEYLAFLNNEPHFGSATVPESVSGFTGDDKYE 237
 Db 175 SATVADFSGTDLILIR--GFLRTERSDLQK--LNNPNVNTMEY-----NDYTF 220
 QY 238 FFEREAVSIDCYAEQVAVARVCCKMGCAFTLORNTTFLKARLACASAP-NMQLYFNQ 296
 Db 221 FFEREAVIEINCKAIVSRVAVCKHDKGPHQFGDRTSLFLSKNSCPDYPFVNE 280
 QY 297 LQAMHTLQDTSMHNT---FPGVFOAQMGDMYLSAICEYOLEIQRVFBGPKEVHEE 353
 Db 281 IQSTSDIIEGNAGGVQEKILGVTFTPVNSIGSVAFCASMSLIESPDGPEDEBTNNS 340
 QY 354 KMDRTD-PVSPRPGSCINNMHRHGTSSLELPDNLNFKVKKHPLMEQVGRW 408
 Db 341 NMLAVPSLAKVPEPRRQGVND-----SKTLDVSYNFKSKITLDEAVPAFTFRI 391
 QY 413 LVKRGF--NFTHLVAD-RVYGLDQATYVLFEGNGDMILKAVSLGPMVH-----LIE 462
 Db 392 LIRISLQYRFTKIAVDQVRRPDGKAYDVLFTGDKVYIKALNSASFSSDPTVDSVYIE 451
 QY 463 ELQLELDQ-EPMSIVL-----SQSKLLFAGSRQVQLVADCIKXKSCADYALADPY 516
 Db 452 ELQVLPQGVPAKMLYVRMDGDSKLVVYSDDEILAKILHRCGSRKITNCRCVSELQDPY 511
 QY 517 CAMSVNTRCYAVAGHFGESL---LIQHWMTSDTSGICNLKSGKIOSGXPKNITVYAGT 572
 Db 512 CAMDVNLTACTVAGSPDMSAGKRRRTIQNISLGEHNAAGCRQGTETIVASVVPPTPTKSSG 571
 QY 573 DLVLPCH 579
 Db 572 DVAHSIH 578

RESULT 15

C49423
 semaphorin II precursor - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 07-May-1999
 C:Accession: C49423
 R:Kolodkin, A.L.; Mathies, D.J.; Goodman, C.S.
 Cell 75, 1389-1399, 1993
 A:Title: The Semaphorin genes encode a family of transmembrane and secreted growth co
 A:Reference number: A49423; MUID:94094332; PMID:8269517
 A:Accession: C49423
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
 A:Molecule type: mRNA
 A:Residues: 1-724 <KOL>
 A:Cross-references: GB:L26083
 C:Genetics:
 A:Gene: sema II
 A:Cross-references: FlyBase:Bgn0011260
 C:Superfamily: semaphorin

Query Match 21.3%; Score 685; DB 2; Length 724;
 Best Local Similarity 32.2%; Pred. No. 2.5e-48;
 Matches 185; Conservative 92; Mismatches 219; Indels 78; Gaps 23;
 QY 42 VRRSQTQIGQDFLTLEPTGLYVGAREALFAFSSEALEL---QGAISWEAPYEKRT 97
 Db 58 VREFN-CGRLLYRFTFMNEDRDTLYVGAMDVFRVNLNISSNCRPAINLEPRDYY 116
 QY 98 ECIQGNKNOTECFNFIRFLQPYN-ASHLYVCGTAFQPKCTYVNMILFTT-LEHGEFEDG 155
 Db 117 SCVSGKSGQIFDCKKHVYKIQSMQGDRLYVCGTNAHNRK-DIVIYANLTHLPRSEIYIG 175
 QY 156 KG---KCPYDPAKCHAGLYDG-----ELYSATLNNFLTETPIILIRNMGPHHSMT-- 202
 Db 176 VGLGIAKQPYDPLDNSTAIYVENGNPGGLPGIYSTNAFTADVITRTDLYNYSARKL 235
 QY 203 -----TEYLAFLNNEPHFGSATVPESVSGFTGDDDKYTFPFRERAVESDCYAEQVYA 255
 Db 236 EYKFRRTIKYDSKWKIDNFGVSF-----DIGEY-----YVFFPRETAVEYINCGKAVYS 285

```
OY 256 RVARVCKGDMGARTLQRRKWTPLKRLACS-APNQLYENQLOAHNTLODTSMHNTPE 314
DB 286 RIAARVCKKDVGGKMLAHNMATYTLKARLNCISGEPFYENEIOYQLPS--DKSRPF 342
OY 315 GVFOAQMGDMYLSAICEYOLFEIORVEGPKYEYHEAQW-DRYTDPVPSPRGSCINN 373
DB 343 ATFTTSTNGLISAVCSFHINEIOAAFNGKFKQSSNSAMLPVLNSRVPFEPFGTCVND 402
OY 374 WHRRHGYTSSLELPDNLNFVKRHPMLMEQVGRMRSPDLLVKRGTFHLYADRY-TGLD 432
DB 403 -----TSN--LPDTVLNFRSHPLMDKAVNHEHNNPYYKRDVFTKLYVDKIRIDL 453
OY 433 GATYTVLFTGTGDCMLKAV-----SLGPNVHLIEIQLFDEPMSLVLSQSKLLP 485
DB 454 NOEYIYYVGTNGRIYKIYOYRNGESLS--KLIDIFEVAPNEAIOYMEISQTRKSLY 510
OY 486 AGSRVOLVLPVADC-IKYRSCADCYLARDPYCAMSNTSRCAVAGHFGSLLIOHMTS 544
DB 511 IGTDRIKOIDIAMCNRRTDNCRCV--RDPYCGMDKEANTC---RPELDLLODV-AN 563
OY 545 DTSGICNLNGSKIOSGPPXPKNITVAVGTDVLVLP 578
DB 564 ETSIDIC-----SSVLKKKIYVYCGSVHLGC 590
```

Search completed: August 7, 2003, 13:50:46
Job time : 46 secs

XX (CURA-) CURAGEN CORP.
 PA Shinkets RA, Fernandes E, Vernet C, Yang M, Boldog FL;
 PI Hermann JL;
 DR WPI, 2001-071385/08.
 DR N-PSDB; AAC84888.
 XX Polynucleotides encoding SECX proteins useful for treating disease
 PT characterized by an aberrant level of cell proliferation and/or
 PT differentiation like cancer or immune associated disorders -
 XX Claim 1; Fig 7; 132pp; English.
 PS
 XX The invention relates to human SECX polypeptides and polynucleotides
 CC encoding them. The SECX polypeptides can be expressed by standard
 CC recombinant methodology. The SECX polypeptides are useful for treating
 CC or preventing a SECX-associated disorder. The invention is useful in
 CC screening assays; detection assays (e.g. chromosomal mapping, cell and
 CC tissue typing, forensic biology); predictive medicine (diagnostic assays,
 CC prognostic assays, monitoring clinical trials, and pharmacogenomics); and
 CC methods of treatment (e.g. therapeutic and prophylactic), especially
 CC disorders characterized by aberrant cell proliferation and/or
 CC differentiation like cancer or immune associated disorders or gestational
 CC disease. The present sequence represents a SEC7 protein.
 CC
 XX
 XX Sequence 596 AA:
 SQ
 Query Match 99.9%; Score 3217; DB 22; Length 596;
 Best Local Similarity 99.8%; Pred. No. 3.7e-317;
 Matches 595; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MAPHNAVMLLAARLMGIGIGAEVWNLVPRKTVSSGELATVVRFSQGTGIDFLITLITE 60
 DB 1 MAPHNAVMLLAARLMGIGIGAEVWNLVPRKTVSSGELATVVRFSQGTGIDFLITLITE 60
 QY 61 PFGILYVGAREALFAFSMEALELOGAISWEAPVEKTECIQKKNNOTECFNFIRLQPY 120
 DB 61 PFGILYVGAREALFAFSMEALELOGAISWEAPVEKTECIQKKNNOTECFNFIRLQPY 120
 QY 121 NASHLYVCGTYAFOPKCTVYNNMLFTLENGEPEDGKCKPRDPKAGHAGLVYDGEYSAT 180
 DB 121 NASHLYVCGTYAFOPKCTVYNNMLFTLENGEPEDGKCKPRDPKAGHAGLVYDGEYSAT 180
 QY 181 LNMFLETEPILIRNMGPHHSMKTEYLAFWLNPEHFVSAYVPEVSFGDDDKVYFFER 240
 DB 181 LNMFLETEPILIRNMGPHHSMKTEYLAFWLNPEHFVSAYVPEVSFGDDDKVYFFER 240
 QY 241 ERAVESDCYAEQVAVARVAVCKDNGGARTLQKKTTPFKARLACGAPPMOLYFNOLQM 300
 DB 241 ERAVESDCYAEQVAVARVAVCKDNGGARTLQKKTTPFKARLACGAPPMOLYFNOLQM 300
 QY 301 HTLDOTSMHNTFFGCVFOAQMGDMYLSACEYQLEIQRVFEQSPRYEHEAOKMRYND 360
 DB 301 HTLDOTSMHNTFFGCVFOAQMGDMYLSACEYQLEIQRVFEQSPRYEHEAOKMRYND 360
 QY 361 PVPSPRPGSCINNMHRHRYGYSLELDPNILNFKKHPMLEQVGRMRSRLPLVKKGTNF 420
 DB 361 PVPSPRPGSCINNMHRHRYGYSLELDPNILNFKKHPMLEQVGRMRSRLPLVKKGTNF 420
 QY 421 THLVADRYTGLDGAITYTLVFIGTGDGMLKAVSLGFWHLIEBLQLPDDEPKSLVLSOS 480
 DB 421 THLVADRYTGLDGAITYTLVFIGTGDGMLKAVSLGFWHLIEBLQLPDDEPKSLVLSOS 480
 QY 481 KKLTFAGRSQSLVOLPVADCIKRYSCADCVLARDPCASVMTSRVAAAGHGSLTIH 540
 DB 481 KKLTFAGRSQSLVOLPVADCIKRYSCADCVLARDPCASVMTSRVAAAGHGSLTIH 540
 QY 541 VMTSDTSGICNLGRSKTQSGPLPKNTTVAGTDVLVPCHLSSNTALPDSNPRESSV 596
 DB 541 VMTSDTSGICNLGRSKTQSGPLPKNTTVAGTDVLVPCHLSSNTALPDSNPRESSV 596

RESULT 2
 AAB48373
 ID AAB48373 standard; Protein; 590 AA.
 XX
 AC AAB48373;
 XX
 DT 20-APR-2001 (first entry)
 XX
 DE Human SEC6 protein sequence (clone ID 20422974.0.133).
 KW SEC6; cytosolic; gynecological; gene therapy; screening assay; human;
 KW SEC6; chromosomal mapping; forensic biology; cell proliferation; cancer;
 KW cell differentiation; immune associated disorder; gestational disease.
 XX
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FT Peptide 1..20
 FT /note- "signal peptide"
 FT Protein 21..590
 FT /note- "mature protein"
 XX
 PX W0200078602-A2.
 XX
 PD 28-DEC-2000.
 XX
 PF 23-JUN-2000; 2000MO-US17328.
 XX
 PR 23-JUN-1999; 990S-0140584.
 PR 20-JUL-1999; 990S-0144722.
 PR 16-SEP-1999; 990S-0154520.
 PR 22-JUN-2000; 2000US-0604286.
 XX
 PA (CURA-) CURAGEN CORP.
 PI Shinkets RA, Fernandes E, Vernet C, Yang M, Boldog FL;
 PI Hermann JL;
 DR WPI, 2001-071385/08.
 DR N-PSDB; AAC84887.
 XX Polynucleotides encoding SECX proteins useful for treating disease
 PT characterized by an aberrant level of cell proliferation and/or
 PT differentiation like cancer or immune associated disorders -
 XX Claim 1; Fig 6; 132pp; English.
 PS
 XX The invention relates to human SECX polypeptides and polynucleotides
 CC encoding them. The SECX polypeptides can be expressed by standard
 CC recombinant methodology. The SECX polypeptides are useful for treating
 CC or preventing a SECX-associated disorder. The invention is useful in
 CC screening assays; detection assays (e.g. chromosomal mapping, cell and
 CC tissue typing, forensic biology); predictive medicine (diagnostic assays,
 CC prognostic assays, monitoring clinical trials, and pharmacogenomics); and
 CC methods of treatment (e.g. therapeutic and prophylactic), especially
 CC disorders characterized by aberrant cell proliferation and/or
 CC differentiation like cancer or immune associated disorders or gestational
 CC disease. The present sequence represents a SEC6 protein.
 CC
 XX
 XX Sequence 590 AA:
 SQ
 Query Match 96.8%; Score 3115.5; DB 22; Length 590;
 Best Local Similarity 99.1%; Pred. No. 7.5e-307;
 Matches 577; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
 QY 1 MAPHNAVMLLAARLMGIGIGAEVWNLVPRKTVSSGELATVVRFSQGTGIDFLITLITE 60
 DB 1 MAPHNAVMLLAARLMGIGIGAEVWNLVPRKTVSSGELATVVRFSQGTGIDFLITLITE 60
 QY 61 PFGILYVGAREALFAFSMEALELOGAISWEAPVEKTECIQKKNNOTECFNFIRLQPY 120
 DB 61 PFGILYVGAREALFAFSMEALELOGAISWEAPVEKTECIQKKNNOTECFNFIRLQPY 120

```

OY 121 NASHLYVCGTAFQPKCTYVNMFLFTLEHGEFEDGKGCPTDPAGHAGLVLDGELYSAT 180
DB 121 NASHLYVCGTAFQPKCTYVNMFLFTLEHGEFEDGKGCPTDPAGHAGLVLDGELYSAT 180
OY 181 LNNFLGTEPIILRNNGPHHSMKTEYLAFWLNPHFVGSAYVPESGFTGDDDKYFFPR 240
DB 181 LNNFLGTEPIILRNNGPHHSMKTEYLAFWLNPHFVGSAYVPESGFTGDDDKYFFPR 240
OY 241 ERAVESDCYAEQVAVARVAVKCGMGARTLQKKTTEFLKARLACAPNMOLYFNQLOAM 300
DB 241 ERAVESDCYAEQVAVARVAVKCGMGARTLQKKTTEFLKARLACAPNMOLYFNQLOAM 300
OY 301 HTLDQTSWNTTFFGVFOAQMGDMYLSAICEYOLEEIORVREGPYKEYHEEAQKMDRYTD 360
DB 301 HTLDQTSWNTTFFGVFOAQMGDMYLSAICEYOLEEIORVREGPYKEYHEEAQKMDRYTD 360
OY 361 PVSPRPGSCINNMHRRHGYSSLELPTNITLNFVKHPLMEQVGPBRMSRPLLYKGTNF 420
DB 361 PVSPRPGSCINNMHRRHGYSSLELPTNITLNFVKHPLMEQVGPBRMSRPLLYKGTNF 420
OY 421 THLVADRYTGIDGATYVLTFTIGTGDGMLKAVSLGPMVHLLELOLFOEPPARSLVLSOS 480
DB 421 THLVADRYTGIDGATYVLTFTIGTGDGMLKAVSLGPMVHLLELOLFOEPPARSLVLSOS 480
OY 481 KLLFAGRSQDLVOLPVADCIKYNACADCVLARDPYCAMSVMTSRCAVAGHFGSLTIOH 540
DB 481 KLLFAGRSQDLVOLPVADCIKYNACADCVLARDPYCAMSVMTSRCAVAGHFGSLTIOH 540
OY 541 VMTSDTSGICMLRGSKIQSGPKNITVYAGTDVLPCHLS 582
DB 541 VMTSDTSGICMLRGSK-KVPRPKNITVYAGTDVLPCHLS 581

```

RESULT 3
AAE03640
ID AAE03640 standard; Protein: 833 AA.

AAE03640;

06-AUG-2001 (first entry)

```

XX Human extracellular matrix and cell adhesion molecule-4 (XMAD-4).
DE
XX
XX Human; extracellular matrix and cell adhesion molecule; XMAD;
XX gene therapy; genetic disorder; adrenoleukodystrophy; leukaemia;
XX Down's syndrome; cystic fibrosis; Gaucher's disease; myotonic dystrophy;
XX sickle cell anaemia; thalassaemia; autoimmune disorder; adenocarcinoma;
XX inflammatory disorder; acquired immune deficiency syndrome; AIDS;
XX Addison's disease; allergy; anaemia; asthma; atherosclerosis; melanoma;
XX Crohn's disease; diabetes mellitus; atopic dermatitis; lymphoma; cancer;
XX glomerulonephritis; multiple sclerosis; Grave's disease; osteoarthritis;
XX osteoporosis; psoriasis; rheumatoid arthritis; ulcerative colitis;
XX infection; cell proliferative disorder; actinic keratosis; myeloma;
XX arteriosclerosis; neotrophic; anticonvulsant; antithyroid; nephrotropic;
XX neuroprotective; dermatological.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..23
FT /label= Signal_peptide
FT Protein 24..833
FT /note= "Mature human extracellular matrix and cell
FT adhesion molecule (XMAD)"
FT Domain 53..481
FT /note= "Semaphorin domain"
XX
XX WO200142285-A2.
XX 14-JUN-2001.
XX PD 05-DEC-2000; 2000WO-US32990.
PP

```

```

XX 10-DEC-1999; 99US-0172852.
PR 16-DEC-1999; 99US-0172354.
XX
XX (INCY-) INCYTE GENOMICS INC.
PI Yue H, Tang YT, Lal P, Burford N, Azimzai Y, Patterson C;
PI Baughn WR, Lu DM, Shah P, Au-Young J;
XX
XX WPI; 2001-381632/40.
DR N-PSDB; AAD08048.
XX
XX New human extracellular matrix and cell adhesion molecules and
XX polynucleotide sequences encoding them, useful for diagnosis,
XX prevention, treatment of genetic, autoimmune and cell proliferative
XX disorders
XX
XX Claim 1; Page 96-98; 135pp; English.
XX
XX The present sequence is a human extracellular matrix and cell
XX adhesion molecule (XMAD). The XMAD is used for screening a compound for
XX effectiveness as an agonist or antagonist of XMAD. The identified agonist
XX or antagonist are used for treating a disease or condition associated
XX with decreased or increased expression of functional XMAD. The
XX polynucleotides encoding XMAD are useful in somatic or germ-line gene
XX therapy to correct a genetic deficiency, to express a conditionally
XX lethal gene product and to express a protein which affords protection
XX against intracellular parasites and also for diagnosis of disorders
XX associated with expression of XMAD. They are also used for generating
XX hybridisation probes useful in mapping the naturally occurring genomic
XX sequences and to create knock in humanised animals (pigs) or transgenic
XX animals (mice or rats) to model human diseases. Oligonucleotide or longer
XX fragments derived from the polynucleotide sequences may be used as
XX elements on a microarray. Antibodies which specifically bind XMAD may be
XX used for the diagnosis of disorders associated with the expression of
XX XMAD, or in assays to monitor patients being treated with XMAD. Diseases
XX diagnosed, prevented or treated include genetic disorders such as
XX adrenoleukodystrophy, Down's syndrome, cystic fibrosis, Gaucher's
XX disease, myotonic dystrophy, sickle cell anaemia, thalassaemia,
XX autoimmune/inflammatory disorders such as acquired immune deficiency
XX syndrome (AIDS), Addison's disease, allergies, anaemia, asthma,
XX atherosclerosis, Crohn's disease, diabetes mellitus, atopic dermatitis,
XX glomerulonephritis, multiple sclerosis, Grave's disease, osteoarthritis,
XX osteoporosis, psoriasis, rheumatoid arthritis, ulcerative colitis,
XX bacterial, fungal, parasitic, protozoal and helminthic infections and
XX cell proliferative disorders such as actinic keratosis, arteriosclerosis
XX and cancer including breast, bladder, bone marrow, brain and uterus
XX cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma.
XX
XX Sequence 833 AA:
SQ
XX
XX Query Match 96.7%; Score 3113.5; DB 22; Length 833;
XX Best Local Similarity 96.3%; Pred. No. 2.1e-306;
XX Matches 579; Conservative 4; Mismatches 7; Indels 11; Gaps 2;
XX
XX 1 MAPHAWMLAARLNGLGIGAEVMMNLVPRKTVSSGELATVYRRSQTGIDPFLTTLRE 60
DB 1 MAPHAWMLAARLNGLGIGAEVMMNLVPRKTVSSGELATVYRRSQTGIDPFLTTLRE 60
XX
XX 61 PFGLLVYGARREALFAFSMALDELQAGISREAVVEKTEIOGKNNQTECFNPIRFLOPY 120
DB 61 PFGLLVYGARREALFAFSMALDELQAGISREAVVEKTEIOGKNNQTECFNPIRFLOPY 120
XX
XX 121 NASHLYVCGTAFQPKCTYVNMFLFTLEHGEFEDGKGCPTDPAGHAGLVLDGELYSAT 180
DB 121 NASHLYVCGTAFQPKCTYVNMFLFTLEHGEFEDGKGCPTDPAGHAGLVLDGELYSAT 180
XX
XX 181 LNNFLGTEPIILRNNGPHHSMKTEYLAFWLNPHFVGSAYVPESGFTGDDDKYFFPR 240
DB 181 LNNFLGTEPIILRNNGPHHSMKTEYLAFWLNPHFVGSAYVPESGFTGDDDKYFFPR 240
XX
XX 241 ERAVESDCYAEQVAVARVAVKCGMGARTLQKKTTEFLKARLACAPNMOLYFNQLOAM 300
DB 241 ERAVESDCYAEQVAVARVAVKCGMGARTLQKKTTEFLKARLACAPNMOLYFNQLOAM 300
XX

```

Db 241 ERALESDCYAEVAVARAVKCGDKMGANTLQKMTTFKLARLACSAFPMQLYFNQLQAM 300
 QY 301 HTLQDTSMHNTTFFGVFOAQMGDMYLSAICEYOLEEIOHFEGPYKEYHEAQRMDRYTD 360
 Db 301 HTLQDTSMHNTTFFGVFOAQMGDMYLSAICEYOLEEIOHFEGPYKEYHEAQRMDRYTD 360
 QY 361 PVPSPRPGSCINNMHRHGYTSSLELPDNLNFVKHHPIMEBOYGRPMSPRLLYKKGTFN 420
 Db 361 PVPSPRPGSCINNMHRHGYTSSLELPDNLNFVKHHPIMEBOYGRPMSPRLLYKKGTFN 420
 QY 421 THLYADRVGTGLDQATYTVLEFIGDGMCLKAVSLGPMVHLIEELOFDEPMRSLVLSQS 480
 Db 421 THLYADRVGTGLDQATYTVLEFIGDGMCLKAVSLGPMVHLIEELOFDEPMRSLVLSQS 480
 QY 481 KLLFAGSRSQVLQVLPVADCIKRYSCADCVLADPFCAMSVNTRCVAAGHRSLLIOH 540
 Db 481 KLLFAGSRSQVLQVLPVADCIKRYSCADCVLADPFCAMSVNTRCVAAGHRSLLIOH 540
 QY 541 VMTSDTSGICNLGRSGTOSGPRPKNTTVVAGTDLVLPCHLSSNLA-----LPDSN 590
 Db 541 VMTSDTSGICNLGRSGTOSGPRPKNTTVVAGTDLVLPCHLSSNLA-----LPDSN 590
 QY 591 P 591
 Db 600 P 600

RESULT 4
 AAB48378
 ID AAB48378 standard; protein; 624 AA.

AC AAB48378;
 DT 20-APR-2001 (first entry)

Human SEC11 protein sequence (clone ID 20422974.0.132-ext2).

SECK: cytostatic; gynecological; gene therapy; screening assay; human;
 SEC11: chromosomal mapping; forensic biology; cell proliferation; cancer;
 cell differentiation; immune associated disorder; gestational disease.

Homo sapiens.

Key Location/Qualifiers
 Peptide 1..20
 Protein /note- "signal peptide"
 21..624
 /note- "mature protein"

MO200078802-A2.

28-DEC-2000.

23-JUN-2000; 2000MO-US17328.

23-JUN-1999; 99US-0140584.

20-JUL-1999; 99US-0144722.

16-SEP-1999; 99US-0154520.

22-JUN-2000; 2000US-0604286.

(CURA-) CURAGEN CORP.

Shinketsu RA, Fernandes E, Vernet C, Yang M, Boldog FL;

Herrmann JI;

WPI; 2001-071385/08.

N-PSDB; AAC84892.

Polynucleotides encoding SECK proteins useful for treating disease

PT characterized by an aberrant level of cell proliferation and/or

PT differentiation like cancer or immune associated disorders -

PS Claim 1; Fig 11; 132pp; English.

XX The invention relates to human SECK polypeptides and polynucleotides
 CC encoding them. The SECK polypeptides can be expressed by standard
 CC recombinant methodology. The SECK polypeptides are useful for treating
 CC or preventing a SECK-associated disorder. The invention is useful in
 CC screening assays; detection assays (e.g. chromosomal mapping, cell and
 CC tissue typing, forensic biology); predictive medicine (diagnostic assays,
 CC prognostic assays, monitoring clinical trials, and pharmacogenomics); and
 CC methods of treatment (e.g. therapeutic and prophylactic), especially
 CC disorders characterized by aberrant cell proliferation and/or
 CC differentiation like cancer or immune associated disorders or gestational
 CC disease. The present sequence represents a SEC11 protein.

Sequence 624 AA;

Query Match 96.5%; Score 3107.5; DB 22; Length 624;
 Best Local Similarity 99.0%; Pred. No. 5.4e-306;

Matches 576; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

1 MAPHMAVWLLAARLMLGIGAGAEYMNLLVPKRTVSSGELATVVRFSQTGIQDFLTTLTE 60

1 MAPHMAVWLLAARLMLGIGAGAEYMNLLVPKRTVSSGELATVVRFSQTGIQDFLTTLTE 60

61 PTGLLYVGAREALFASMEALELQAI SWEAPYEKTECIQKRNKNOTECFNRIRLOPY 120

61 PTGLLYVGAREALFASMEALELQAI SWEAPYEKTECIQKRNKNOTECFNRIRLOPY 120

61 PTGLLYVGAREALFASMEALELQAI SWEAPYEKTECIQKRNKNOTECFNRIRLOPY 120

121 NASHLVYCGTYAPOPCTYVNMULTLEHGEFEDGKCPYDPAKGHAGLLVYGEIYSAT 180

121 NASHLVYCGTYAPOPCTYVNMULTLEHGEFEDGKCPYDPAKGHAGLLVYGEIYSAT 180

121 NASHLVYCGTYAPOPCTYVNMULTLEHGEFEDGKCPYDPAKGHAGLLVYGEIYSAT 180

181 LNNFLGTEPIILNNKMPHHSMTETLAFWLNEDPHFVSAYVPSVSGFTGDDDKYVFFFR 240

181 LNNFLGTEPIILNNKMPHHSMTETLAFWLNEDPHFVSAYVPSVSGFTGDDDKYVFFFR 240

181 LNNFLGTEPIILNNKMPHHSMTETLAFWLNEDPHFVSAYVPSVSGFTGDDDKYVFFFR 240

241 ERAVESDCYAEVAVARAVKCGDKMGARTLQKMTTFKLARLACSAFPMQLYFNQLQAM 300

241 ERAVESDCYAEVAVARAVKCGDKMGARTLQKMTTFKLARLACSAFPMQLYFNQLQAM 300

241 ERAVESDCYAEVAVARAVKCGDKMGARTLQKMTTFKLARLACSAFPMQLYFNQLQAM 300

301 HTLQDTSMHNTTFFGVFOAQMGDMYLSAICEYOLEEIOHFEGPYKEYHEAQRMDRYTD 360

301 HTLQDTSMHNTTFFGVFOAQMGDMYLSAICEYOLEEIOHFEGPYKEYHEAQRMDRYTD 360

301 HTLQDTSMHNTTFFGVFOAQMGDMYLSAICEYOLEEIOHFEGPYKEYHEAQRMDRYTD 360

361 PVPSPRPGSCINNMHRHGYTSSLELPDNLNFVKHHPIMEBOYGRPMSPRLLYKKGTFN 420

361 PVPSPRPGSCINNMHRHGYTSSLELPDNLNFVKHHPIMEBOYGRPMSPRLLYKKGTFN 420

361 PVPSPRPGSCINNMHRHGYTSSLELPDNLNFVKHHPIMEBOYGRPMSPRLLYKKGTFN 420

421 THLYADRVGTGLDQATYTVLEFIGDGMCLKAVSLGPMVHLIEELOFDEPMRSLVLSQS 480

421 THLYADRVGTGLDQATYTVLEFIGDGMCLKAVSLGPMVHLIEELOFDEPMRSLVLSQS 480

421 THLYADRVGTGLDQATYTVLEFIGDGMCLKAVSLGPMVHLIEELOFDEPMRSLVLSQS 480

481 KLLFAGSRSQVLQVLPVADCIKRYSCADCVLADPFCAMSVNTRCVAAGHRSLLIOH 540

481 KLLFAGSRSQVLQVLPVADCIKRYSCADCVLADPFCAMSVNTRCVAAGHRSLLIOH 540

481 KLLFAGSRSQVLQVLPVADCIKRYSCADCVLADPFCAMSVNTRCVAAGHRSLLIOH 540

541 VMTSDTSGICNLGRSGTOSGPRPKNTTVVAGTDLVLPCHLSS 582

541 VMTSDTSGICNLGRSGTOSGPRPKNTTVVAGTDLVLPCHLSS 582

RESULT 5

AAE18213

ID AAE18213 standard; protein; 833 AA.

AAE18213;

07-MAY-2002 (first entry)

Human MOL5a protein.

DE Secreted molecule; MOL5a protein; MOLX; cardiomyopathy; atherosclerosis;

XX diabetes; chromosomal disorder; anemia; psoriasis; scarring;

KW liver cirrhosis; neurodegeneration; osteoarthritis; organ rejection;

KM cerebral thrombosis; hypertension; systemic lupus erythematosus; asthma;
 KM immune disease; ischemia; immunodeficiency; rheumatoid arthritis; ulcer;
 KM HIV; human immuno deficiency virus; hepatitis; haemostatic disease; pain;
 KM haematopoietic; thrombolytic; thrombocytopenia; Alzheimer's disease;
 KM Parkinson's disease; Huntington's disease; muscular disease; stress;
 KM ocular disease; growth disorder; depression; epilepsy; contraceptive;
 KM vulnery; osteopathic; haemostatic; tranquilliser; antidepressant;
 KM analgesic; vasotropic; hypotensive; gene therapy; chromosome 2.
 XX Homo sapiens.
 OS XX
 FH Key Location/Qualifiers
 FT Misc-difference 247 /note- "Encoded by GCC"
 FT XX
 XX W0200206339-A2.
 PD 24-JAN-2002.
 XX 03-JUL-2001: 2001WO-US21249.
 PF 03-JUL-2000: 2000US-215854P.
 XX 03-JUL-2000: 2000US-215855P.
 PR 03-JUL-2000: 2000US-215902P.
 PR 07-JUL-2000: 2000US-215855P.
 PR 07-JUL-2000: 2000US-215856P.
 PR 07-JUL-2000: 2000US-216722P.
 PR 17-JUL-2000: 2000US-216622P.
 PR 17-JUL-2000: 2000US-218992P.
 PR 27-JUL-2000: 2000US-221285P.
 PR 14-FEB-2001: 2001US-268734P.
 XX (CURA-) CURAGEN CORP.
 XX Spaderina SK, Tchernev V, Liu X, Shenoy S, Spytek K, zerhusen B;
 PI Patreusjan M, Taupier RJ, Rastelli L, Grose KM, Szekeres ES;
 PI Alsobrook J, Lepley DM, Shen L, Burgess CE, Shinkets RA;
 PI Padigaru M;
 XX WPI: 2002-155038/20.
 DR N-PSDB: AAD28947.
 XX Nucleic acids encoding secreted polypeptides, designated MOLX
 PT polypeptides, useful for treating a MOLX-associated disorder, e.g.
 PT cardiomyopathy, atherosclerosis, diabetes and metabolic disorders -
 XX
 PS Claim 1, Page 60; 223pp: English.
 XX The patent discloses nucleic acid sequences encoding novel secreted
 CC molecule (MOL) polypeptides, designated MOLX polypeptides (i.e. a MOL
 CC protein where X is an integer from 1 to 8). Sequences of the invention
 CC are useful for treating or preventing a MOLX-associated disorder in
 CC humans. They are useful for treating or preventing cardiomyopathy,
 CC atherosclerosis and disorders related to cell signal processing and
 CC metabolic pathway modulation. The MOLX antibodies are useful for
 CC treating or preventing diabetes and disorders related to cell signal
 CC processing and metabolic pathway modulation. MOLX sequences are useful
 CC for the treatment or diagnosis of other MOLX-associated disorders, e.g.
 CC chromosomal disorders, albinism, anemia, liver cirrhosis, psoriasis,
 CC scarring, neurodegeneration, osteoarthritis, organ rejection, cerebral
 CC thrombosis, ischemia, hypertension, systemic lupus erythematosus, immune
 CC diseases, immunodeficiency, HIV (human immuno deficiency virus), viral,
 CC bacterial and fungal infections, hepatitis, rheumatoid arthritis, asthma,
 CC haematopoietic, thrombolytic, haemostatic diseases, thrombocytopenia,
 CC ulcers, Alzheimer's disease, Parkinson's disease, Huntington's disease,
 CC ocular disease, muscular diseases, growth disorders, loss of libido,
 CC stress, depression, pain and epilepsy. They are useful for preventing
 CC chemotherapy side effects and as contraceptives. Sequences of the
 CC invention are also useful for gene therapy. The present sequence
 CC is human semaphorin 4C-like protein, MO15a. MO15a gene is localised
 CC on chromosome 2.
 XX Sequence 833 AA;
 SO

Query Match 96.0%; Score 3089.5; DB 23; Length 833;
 Best Local Similarity 95.7%; Pred. No. 5.8e-304;
 Matches 575; Conservative 4; Mismatches 11; Indels 11; Gaps 2;
 QY 1 MAPHAAVLLAARLMGIGAEVMMNLVPRRTVSSGELATVVRNSQTGTIDFTLTITE 60
 DB 1 MAPHAAVLLAARLMGIGAEVMMNLVPRRTVSSGELATVVRNSQTGTIDFTLTITE 60
 QY 61 PTGLIYVGAARALAFSNEALELOGAISWEAPVEKTECIQKGNNOECENFTRFDLPY 120
 DB 61 PTGLIYVGAARALAFSNEALELOGAISWEAPVEKTECIQKGNNOECENFTRFDLPY 120
 QY 121 NASHLYVCGTYAFQPKCTYNNMLFTTLEHGEFEDKKGKCPYDPAKGAGLIVDGLYSAT 180
 DB 121 NASHLYVCGTYAFQPKCTYNNMLFTTLEHGEFEDKKGKCPYDPAKGAGLIVDGLYSAT 180
 QY 181 LNNTLGTETPILRNNGPHHSKTEYLAFWLNPHFVGSAYVPEVSGSTGDDKRYFFER 240
 DB 181 LNNTLGTETPILRNNGPHHSKTEYLAFWLNPHFVGSAYVPEVSGSTGDDKRYFFER 240
 QY 241 ERAVSDCYAEOYAVARAVCKGDMGARTLQKRTTEFLKARLACSAANMOLYFNQLOAM 300
 DB 241 ERAVSDCYAEOYAVARAVCKGDMGARTLQKRTTEFLKARLACSAANMOLYFNQLOAM 300
 QY 301 HTLQDTSMHNTTFEGVFOAGMDMTLSACEQLEIQRVEGPKYKEHEEAKMDRYTD 360
 DB 301 HTLQDTSMHNTTFEGVFOAGMDMTLSACEQLEIQRVEGPKYKEHEEAKMDRYTD 360
 QY 361 PVPSPRPSCINNMWRHRYGYSLELNDNLNFKKHPDLMEDQVPRNSRPLLYKGTNF 420
 DB 361 PVPSPRPSCINNMWRHRYGYSLELNDNLNFKKHPDLMEDQVPRNSRPLLYKGTNF 420
 QY 421 THIVADRYTGIDGATYTLFTIGDGWLLKAVSGCPWHLLEQLFQEPHRSYLVSQS 480
 DB 421 THIVADRYTGIDGATYTLFTIGDGWLLKAVSGCPWHLLEQLFQEPHRSYLVSQS 480
 QY 481 KKLTFAGRSRDLVOLPVADCIKRRSCADCVLARDPYCAMS VYTSRCVAVGGHFGSLIQH 540
 DB 481 KKLTFAGRSRDLVOLPVADCIKRRSCADCVLARDPYCAMS VYTSRCVAVGGHFGSLIQH 540
 QY 541 VMTSDTSGICNLNRSKIGSGPKNITVAGTDLVLPCHLSNLA-----LPDSN 590
 DB 541 VMTSDTSGICNLNRSKIGSGPKNITVAGTDLVLPCHLSNLAHAMWTEFGGRDLPAHQ 599
 QY 591 P 591
 DB 600 P 600
 DE Human MOL5b protein.
 XX Secreted molecule; MOL5b protein; MOLX; cardiomyopathy; atherosclerosis;
 KM diabetes; chromosomal disorder; albinism; anemia; psoriasis; scarring;
 KM liver cirrhosis; neurodegeneration; osteoarthritis; organ rejection;
 KM cerebral thrombosis; hypertension; systemic lupus erythematosus; asthma;
 KM immune disease; ischemia; immunodeficiency; rheumatoid arthritis; ulcer;
 KM HIV; human immuno deficiency virus; hepatitis; haemostatic disease; pain;
 KM haematopoietic; thrombolytic; thrombocytopenia; Alzheimer's disease;
 KM Parkinson's disease; Huntington's disease; muscular disease; stress;
 KM ocular disease; growth disorder; depression; epilepsy; contraceptive;
 KM vulnery; osteopathic; haemostatic; tranquilliser; antidepressant;
 KM analgesic; vasotropic; hypotensive; gene therapy; chromosome 2.
 XX Homo sapiens.
 OS

XX XX MO200206339-A2.
 XX PD 24-JAN-2002.
 XX PF 03-JUL-2001; 2001MO-US21249.
 XX 03-JUL-2000; 2000US-215854P.
 PR 03-JUL-2000; 2000US-215856P.
 PR 03-JUL-2000; 2000US-215902P.
 PR 07-JUL-2000; 2000US-216585P.
 PR 07-JUL-2000; 2000US-216586P.
 PR 07-JUL-2000; 2000US-216722P.
 PR 17-JUL-2000; 2000US-218642P.
 PR 17-JUL-2000; 2000US-218992P.
 PR 27-JUL-2000; 2000US-221285P.
 PR 14-FEB-2001; 2001US-268734P.
 XX (CURA-) CURAGEN CORP.
 PA Spaderna SR, Tchernev V, Liu X, Shenoy S, Spytek K, Zehrhusen B;
 PI Paturajan M, Taupier RJ, Rastelli L, Grosse WM, Sekeres ES;
 PI Alsobrook J, Lepley DM, Shen L, Burgess CE, Shinkets RA;
 PI Padigaru M;
 DR WPI; 2002-155038/20.
 DR N-PSDB; AAD28948.
 XX Nucleic acids encoding secreted polypeptides, designated MOLX
 PT polypeptides, useful for treating a MOLX-associated disorder, e.g.
 PT cardiomyopathy, atherosclerosis, diabetes and metabolic disorders -
 XX claim 1; Page 61-62; 223pp; English.
 XX The patent discloses nucleic acid sequences encoding novel secreted
 CC molecule (MOL) polypeptides, designated MOLX polypeptides (i.e. a MOL
 CC protein where X is an integer from 1 to 8). Sequences of the invention
 CC are useful for treating or preventing a MOLX-associated disorder in
 CC humans. They are useful for treating or preventing cardiomyopathy,
 CC atherosclerosis and disorders related to cell signal processing and
 CC metabolic pathway modulation. The MOLX antibodies are useful for
 CC treating or preventing diabetes and disorders related to cell signal
 CC processing and metabolic pathway modulation. MOLX sequences are useful
 CC for the treatment or diagnosis of other MOLX-associated disorders, e.g.
 CC chromosomal disorders, albinism, anaemia, liver cirrhosis, psoriasis,
 CC scarring, neurodegeneration, osteoarthritis, organ rejection, cerebral
 CC thrombosis, ischaemia, hypertension, systemic lupus erythematosus, immune
 CC diseases, immunodeficiency, HIV (human immunodeficiency virus), viral,
 CC bacterial and fungal infections, hepatitis, rheumatoid arthritis, asthma,
 CC haematopoietic, thrombolytic, haemostatic diseases, thrombocytopenia,
 CC ulcers, Alzheimer's disease, Parkinson's disease, Huntington's disease,
 CC ocular disease, muscular diseases, growth disorders, loss of libido,
 CC stress, depression, pain and epilepsy. They are useful for preventing
 CC chemotherapy side effects and as contraceptives. Sequences of the
 CC invention are also useful for gene therapy. The present sequence
 CC is human semaphorin 4C-like protein, MOL5b. MOL5b gene is localised
 CC on chromosome 2.
 XX SQ Sequence 833 AA;

Query Match 94.9%; Score 3055.5; DB 23; Length 833;
 .Best Local Similarity 94.7%; Pred. No. 1.7e-300;
 Matches 569; Conservative 4; Mismatches 17; Indels 11; Gaps 2;
 QY 1 MAPHNAVMLLAARLMGIGAGAEVWNLVPRKTVSSGELATVYRFSQGIQDFLTTLTLE 60
 DB 1 MAPHNAVMLLAARLMGIGAGAEVWNLVPRKTVSSGELATVYRFSQGIQDFLTTLTLE 60
 QY 61 PTGLLYVGAREALFAFSEMELELOGAIISWEAPVEKTECIQKGNQTECFNFIREFLOPY 120
 DB 61 PTGLLYVGAREALFAFSEMELELOGAIISWEAPVEKTECIQKGNQTECFNFIREFLOPY 120
 QY 121 NASHLYVCGTYAFQPRCTYVNMILFTLEHGEFEDGKCKCPYDPAKHAGLAVDELYSAT 180
 DB 121 NASHLYVCGTYAFQPRCTYVNMILFTLEHGEFEDGKCKCPYDPAKHAGLAVDELYSAT 180
 QY 181 LNNEFLGTEPTILNNMGPHHSMKTEYLAFMLEHPFVSAVPPSVGSFTGDDDKVYFFER 240
 DB 181 LNNEFLGTEPTILNNMGPHHSMKTEYLAFMLEHPFVSAVPPSVGSFTGDDDKVYFFER 240
 QY 241 ERAVESDCYAEQVYAVARVCKGDMGARTLQKRWTFLLKARLACAPNNQLYFNLOAM 300
 DB 241 ERAVESDCYAEQVYAVARVCKGDMGARTLQKRWTFLLKARLACAPNNQLYFNLOAM 300
 QY 301 HTLQDTSMNHTTFFGVFOAGMDMYLSAICEYOLBEIQRVFBGPKYEHEBAQKWRD 360
 DB 301 HTLQDTSMNHTTFFGVFOAGMDMYLSAICEYOLBEIQRVFBGPKYEHEBAQKWRD 360
 QY 361 PVSPRPGSCINMHRHGGTSLELPDNLNVKHPMLEBOVGPRMSPRLLYKKGDTN 420
 DB 361 PVSPRPGSCINMHRHGGTSLELPDNLNVKHPMLEBOVGPRMSPRLLYKKGDTN 420
 QY 421 THLVADRVTLGDATYTVLFTIGDCWMLKAVSLGPMVHLIEELQFPDEPNRSLVLSOS 480
 DB 421 THLVADRVTLGDATYTVLFTIGDCWMLKAVSLGPMVHLIEELQFPDEPNRSLVLSOS 480
 QY 481 KRLPFGNSQSLQVLPVADCIKTRSCADCVLADPPICAMSNTSRGVAGGHRGSLILIOH 540
 DB 481 KRLPFGNSQSLQVLPVADCIKTRSCADCVLADPPICAMSNTSRGVAGGHRGSLILIOH 540
 QY 541 VMTSDTSIGCNLNGSKR-KVRPTKNTIVAGTDLVPLCHLSSNLAHAKRTFGGRDIPAEQ 599
 DB 541 VMTSDTSIGCNLNGSKR-KVRPTKNTIVAGTDLVPLCHLSSNLAHAKRTFGGRDIPAEQ 599
 QY 591 P 591
 DB 600 P 600
 RESULT 7
 AAE18215 standard; Protein; 833 AA.
 AC AAE18215;
 DE 07-MAY-2002 (first entry)
 DT Human MOL5c protein.
 XX Secreted molecule; MOL5c protein; MOLX; cardiomyopathy; atherosclerosis;
 KW diabetes; chromosomal disorder; albinism; anaemia; psoriasis; scarring;
 KW liver cirrhosis; neurodegeneration; osteoarthritis; organ rejection;
 KW cerebral thrombosis; hypertension; systemic lupus erythematosus; asthma;
 KW immune disease; ischaemia; immunodeficiency; rheumatoid arthritis; ulcer;
 KW HIV; human immunodeficiency virus; hepatitis; haemostatic disease; pain;
 KW haematopoietic; thrombolytic; thrombocytopenia; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; muscular disease; stress;
 KW ocular disease; growth disorder; depression; epilepsy; contraceptive;
 KW vulnary; osteopathic; haemostatic; tranquiliser; antidepressant;
 KW analgesic; vasotropic; hypotensive; gene therapy; chromosome 2.
 XX Homo sapiens.
 OS Homo sapiens.
 PH Key Location/Qualifiers
 FT Peptide 1..20
 FT /label= signal_peptide
 FT Protein 21..833
 FT /label= Mature_MOL5c-protein
 XX MO200206339-A2.
 XX 24-JAN-2002.
 XX 03-JUL-2001; 2001MO-US21249.
 XX 03-JUL-2000; 2000US-215854P.

PR 03-JUL-2000; 2000US-215856P.
PR 03-JUL-2000; 2000US-215802P.
PR 07-JUL-2000; 2000US-216585P.
PR 07-JUL-2000; 2000US-216586P.
PR 17-JUL-2000; 2000US-216722P.
PR 17-JUL-2000; 2000US-218623P.
PR 17-JUL-2000; 2000US-218992P.
PR 27-JUL-2000; 2000US-221285P.
PR 14-FEB-2001; 2001US-268734P.
XX
XX (CURA-) CURAGEN CORP.
XX
PI Spaderna SK, Tchernev V, Liu X, Shenoy S, Spytek K, Zernhusen B;
PI Paturajan M, Taupier RJ, Rastelli L, Grosse WM, Szekeles ES;
PI Alsobrook J, Lepley DM, Shen L, Burgess CE, Shinkets RA;
PI Padigar M;
XX
XX WPI: 2002-155038/20.
DR N-PSDB: AAD28949.
XX
XX Nucleic acids encoding secreted polypeptides, designated MOX
PT polypeptides, useful for treating a MOX-associated disorder, e.g.
PT cardiomyopathy, atherosclerosis, diabetes and metabolic disorders -
XX
XX Claim 1: Page 64; 223pp: English.
XX
XX The patent discloses nucleic acid sequences encoding novel secreted
CC molecule (MOX) polypeptides, designated MOX polypeptides (i.e. a MOX
CC protein where X is an integer from 1 to 8). Sequences of the invention
CC are useful for treating or preventing a MOX-associated disorder in
CC humans. They are useful for treating or preventing cardiomyopathy,
CC atherosclerosis and disorders related to cell signal processing and
CC metabolic pathway modulation. The MOX antibodies are useful for
CC treating or preventing diabetes and disorders related to cell signal
CC processing and metabolic pathway modulation. MOX sequences are useful
CC for the treatment or diagnosis of other MOX-associated disorders, e.g.
CC chromosomal disorders, albinism, anaemia, liver cirrhosis, psoriasis,
CC scarring, neurodegeneration, osteoarthritis, organ rejection, cerebral
CC thrombosis, ischaemia, hypertension, systemic lupus erythematosus, immune
CC diseases, immunodeficiency, HIV (human immunodeficiency virus), viral,
CC bacterial and fungal infections, hepatitis, rheumatoid arthritis, asthma,
CC haemolytic, thrombolytic, haemostatic diseases, thrombocytopaenia,
CC ulcers, Alzheimer's disease, Parkinson's disease, Huntington's disease,
CC ocular disease, muscular diseases, growth disorders, loss of libido,
CC stress, depression, pain and epilepsy. They are useful for preventing
CC chemotherapy side effects and as contraceptives. Sequences of the
CC invention are also useful for gene therapy. The present sequence
CC is human semaphorin 4C-like protein, MOX5c. MOX5c gene is localised
CC on chromosome 2.
XX
XX
SQ Sequence 833 AA;
Query Match 94.9%; Score 3055.5; DB 23; Length 833;
Best Local Similarity 94.7%; Pred. No. 1.7e-300;
Matches 569; Conservative 4; Mismatches 17; Indels 11; Gaps 2;

DB 241 ERAVESDCTAEQVAVARAVARVCKGDMGARTLQRTKMTFFLKARLACSAAPNMOLYFNQLOAM 300
|||||
DB 241 ERAVESDCTAEQVAVARAVARVCKGDMGARTLQRTKMTFFLKARLACSAAPNMOLYFNQLOAM 300
|||||
QY HTLADTSWNTTFFGVFOAGNDYLSAICYOLEIQRVEGPKYKEHEEAKMDRYTD 360
|||||
DB 301 HTLADTSWNTTFFGVFOAGNDYLSAICYOLEIQRVEGPKYKEHEEAKMDRYTD 360
|||||
QY 361 PVSPRPSCINNNHRRHGYTSSLEPDNINLNFYKHPLEBOYGPKNRPLVYKGTNF 420
|||||
DB 361 PVSPRPSCINNNHRRHGYTSSLEPDNINLNFYKHPLEBOYGPKNRPLVYKGTNF 420
|||||
QY 421 THLVADRTGTGAGATTYYLFTGTGDLTKAVSICGPNVHLIEEQLDQEPKRSIVLSQS 480
|||||
DB 421 THLVADRTGTGAGATTYYLFTGTGDLTKAVSICGPNVHLIEEQLDQEPKRSIVLSQS 480
|||||
QY 481 KKLIFAGRSQVLCVPVADCKIKYRSCADCVIARDPYCAWSVNTSCVAVGSHGSLIQH 540
|||||
DB 481 KKLIFAGRSQVLCVPVADCKIKYRSCADCVIARDPYCAWSVNTSCVAVGSHGSLIQH 540
|||||
QY 541 VMTSDTSGICNLKRSKIOSGPKNTVYVACTDVLPCHLSSNLA-----LPDSN 590
|||||
DB 541 VMTSDTSGICNLKRSK-KVRPTPKNTVYVACTDVLPCHLSSNLAHAKWTRGGRDLPAED 599
|||||
QY 591 P 591
DB 600 P 600
|||||
RESULT 8
ABB98402
ID ABB98402 standard; Protein; 805 AA.
XX
XX ABB98402;
AC
AC 21-OCT-2002 (first entry)
XX
XX Human NOV2, a semaphorin like protein.
DE
XX
XX Human NOV2; cytosolic; Cardiac; Antiinflammatory; Immunosuppressive;
KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
KW Antislathermic; Nephrotropic; Hepatotropic; Neutroprotective; Nootropic;
KW Antibacterial; Viaricide; Antiparasitic; Relaxant; Anticonvulsant;
KW Gene therapy; NOV; cancer; heart disease; inflammation;
KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes;
KW obesity; asthma; Iga nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; infection; stroke; muscular dystrophy; epilepsy;
KW wasting disorder; semaphorin-like protein.
XX
XX Homo sapiens.
OS
XX
XX WO20025704-A2.
PN
XX
XX 18-JUL-2002.
PD
XX
XX 09-JAN-2002; 2002WO-US00554.
PF
XX
XX 09-JAN-2001; 2001US-260417P.
PR 10-JAN-2001; 2001US-260831P.
PR 28-FEB-2001; 2001US-272338P.
PR 09-MAR-2001; 2001US-274876P.
PR 18-APR-2001; 2001US-284704P.
PA
XX
XX (CURA-) CURAGEN CORP.
XX
PI Padigar M, Li L, Zernhusen BD, Casman SJ, Shenoy S, Spytek KA;
PI Zhong M, Gangolli EA, Burgess CE, Paturajan M, Vernet CAM;
PI Taylor S, Tchernev V, Miller CE, Guo X, Boldog FL, Grosse WM;
PI Alsobrook J, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;
PI MacDougall J, Malysankar U, Millet I, Peyman J, Smithson G;
PI Gunther E, Stone DJ;
XX
XX WPI: 2002-590674/63.
DR N-PSDB: ABB985379.

XX NOVX polypeptides and encoding polynucleotides, useful for preventing
PT or treating NOVX-associated disorders e.g. cancer, inflammation, or
PT Alzheimer's disease, and in chromosome mapping, tissue typing or
PT pharmacogenomics

PS Claim 1, Page 20; 358pp; English.

CC The present sequence is the protein sequence for a NOV protein. The
CC NOV proteins and coding sequences are useful for treating or preventing
CC NOV-associated disorders or in the manufacture of a medicament for
CC treating the disorders, such as cancer, heart disease, inflammation,
CC autoimmune disorders, allergies, blood disorders, AIDS, diabetes,
CC obesity, asthma, Iga nephropathy, cirrhosis, arthritis, Alzheimer's
CC disease, infections (e.g. bacterial, viral, parasitic), stroke, muscular
CC dystrophy, epilepsy, and other wasting disorders associated with chronic
CC diseases. NOV2 is a semaphorin like protein.

XX Sequence 805 AA;

Query Match 90.9%; Score 2927.5; DB 23; Length 805;
Best Local Similarity 91.7%; Pred. No. 1.6e-287;
Matches 551; Conservative 2; Mismatches 9; Indels 39; Gaps 3;

OY 1 MAPHNAVWLLAARLNGIGAEVWNNLVPRKTVSSGELATVYRRSQTGIODFLTITLRE 60
DB 1 MAPHNAVWLLAARLNGIGAEVWNNLVPRKTVSSGELATVYRRSQTGIODFLTITLRE 60
OY 61 PTGLLYVGAREALFAFMSMALELOGAIISWEAPVEKTECIQKNNQTCFNFIRLOPY 120
DB 61 PTGLLYVGAREALFAFMSMALELOGAIISWEAPVEKTECIQKNNQTCFNFIRLOPY 120
OY 121 NASHLYVCGTVAFOPKCTVYNNMLFTLEHGEFEDGKGCPRYPANAGHLLVDGELYSAT 180
DB 121 NASHLYVCGTVAFOPKCTVYNNMLFTLEHGEFEDGKGCPRYPANAGHLLVDGELYSAT 180
OY 181 LNPFLETEPIILRNKGPBHSMTETYLAFWLNPHFVGSAYVESVGSFSGDDDKYFFER 240
DB 181 LNPFLETEPIILRNKGPBHSMTETYLAFWLNPHFVGSAYVESVGSFSGDDDKYFFER 240
OY 241 ERAVESDCYAEQVAVARAVARVCKDGMGARTLQKMTTFELKARLACAPMMQLYFNLOLM 300
DB 241 ERAVESDCYAEQVAVARAVARVCKDGMGARTLQKMTTFELKARLACAPMMQLYFNLOLM 300
OY 213 ERAVESDCYAEQVAVARAVARVCKDGMGARTLQKMTTFELKARLACAPMMQLYFNLOLM 272
DB 213 ERAVESDCYAEQVAVARAVARVCKDGMGARTLQKMTTFELKARLACAPMMQLYFNLOLM 272
OY 301 HTLODTSMHNTTFFGVFOAQMGDMYLSAICEYOLEEIQVFEGPYKEVHEEAQKMDRYD 360
DB 301 HTLODTSMHNTTFFGVFOAQMGDMYLSAICEYOLEEIQVFEGPYKEVHEEAQKMDRYD 360
OY 273 HTLODTSMHNTTFFGVFOAQMGDMYLSAICEYOLEEIQVFEGPYKEVHEEAQKMDRYD 332
DB 273 HTLODTSMHNTTFFGVFOAQMGDMYLSAICEYOLEEIQVFEGPYKEVHEEAQKMDRYD 332
OY 361 PVPSPRPGSCINNMHRRHGYTSSLELPONILNFVKKHPIMEBOVGRMSRPLLYKKGTFN 420
DB 361 PVPSPRPGSCINNMHRRHGYTSSLELPONILNFVKKHPIMEBOVGRMSRPLLYKKGTFN 420
OY 333 PVPSPRPGSCINNMHRRHGYTSSLELPONILNFVKKHPIMEBOVGRMSRPLLYKKGTFN 392
DB 333 PVPSPRPGSCINNMHRRHGYTSSLELPONILNFVKKHPIMEBOVGRMSRPLLYKKGTFN 392
OY 421 THLVADRYTGLDGAATYTVLFTIGTGDMWLKAVSLSPVHLLIELOLFDQEPKSLVLSOS 480
DB 421 THLVADRYTGLDGAATYTVLFTIGTGDMWLKAVSLSPVHLLIELOLFDQEPKSLVLSOS 480
OY 393 THLVADRYTGLDGAATYTVLFTIGTGDMWLKAVSLSPVHLLIELOLFDQEPKSLVLSOS 452
DB 393 THLVADRYTGLDGAATYTVLFTIGTGDMWLKAVSLSPVHLLIELOLFDQEPKSLVLSOS 452
OY 481 KLLLEAGRSQSLVQLPVADCIYRSCADCVLARDPYCAMSVTSRCVAVAGHSGSLIIOH 540
DB 481 KLLLEAGRSQSLVQLPVADCIYRSCADCVLARDPYCAMSVTSRCVAVAGHSGSLIIOH 540
OY 453 KLLLEAGRSQSLVQLPVADCIYRSCADCVLARDPYCAMSVTSRCVAVAGHSGSLIIOH 512
DB 453 KLLLEAGRSQSLVQLPVADCIYRSCADCVLARDPYCAMSVTSRCVAVAGHSGSLIIOH 512
OY 541 VMTSDTSGICNLGSKIGSKIPKNTVYAGTDVLPCHLSSNLA-----LPDSN 590
DB 541 VMTSDTSGICNLGSKIGSKIPKNTVYAGTDVLPCHLSSNLA-----LPDSN 590
OY 591 P 591
DB 591 P 591
OY 572 P 572
DB 572 P 572

RESULT 9
ABB97963
ID ABB97963 standard; Protein; 886 AA.
XX

AC ABB97963;
XX
XX 06-SEP-2002 (first entry)
XX
XX Human protein sequence #30.

XX Human; brain; tonsil; hippocampus; foetal brain; diagnosis.

XX Homo sapiens.

XX MO200252005-A1.

XX 04-JUL-2002.

XX 20-DEC-2001; 2001WO-JP11217.

XX 22-DEC-2000; 2000JP-0389742.

XX (KAZU-) KAZUSA DNA RES INST FOUND.

XX (CELE-) CELESTAR LEXICO-SCI LTD.

XX Ohara O, Nagase T, Nakajima D;

XX WPI, 2002-500762/53.

XX N-PSDB; ABN83983.

XX Claim 1(a); Page 119-125; 238pp; Japanese.

CC The invention relates to DNA encoding polypeptides directly cloned from
CC cDNA libraries originating in adult whole brain, human tonsil, human
CC adult hippocampus and human foetal whole brain. Polypeptides and
CC polynucleotides of the invention may be used in the investigation of
CC differential expression of the DNA sequences in normal subjects and
CC disease patients. They may also be used in the production of antibodies,
CC oligonucleotide probes and DNA chips for diagnosis and identification
CC of drugs for treatment of diseases with which the DNA sequences are
CC associated. The sequences given in records ABB97934-ABB97964 represent
CC human proteins of the invention.

XX Sequence 886 AA;

Query Match 90.6%; Score 2918; DB 23; Length 886;
Best Local Similarity 92.2%; Pred. No. 1.7e-286;
Matches 554; Conservative 5; Mismatches 22; Indels 20; Gaps 5;

OY 8 W-LIARLNGIGAEVWNNLVPRKTVSS-----GELATVYRRSQTGIODFLTITLRE 60
DB 8 W-LIARLNGIGAEVWNNLVPRKTVSS-----GELATVYRRSQTGIODFLTITLRE 60
OY 56 WHLVAV--WGAGSRGARLRAVEPOGCPSPAMLTLPALATVYRRSQTGIODFLTITLRE 113
DB 56 WHLVAV--WGAGSRGARLRAVEPOGCPSPAMLTLPALATVYRRSQTGIODFLTITLRE 113
OY 61 PTGLLYVGAREALFAFMSMALELOGAIISWEAPVEKTECIQKNNQTCFNFIRLOPY 120
DB 61 PTGLLYVGAREALFAFMSMALELOGAIISWEAPVEKTECIQKNNQTCFNFIRLOPY 120
OY 114 PTGLLYVGAREALFAFMSMALELOGAIISWEAPVEKTECIQKNNQTCFNFIRLOPY 173
DB 114 PTGLLYVGAREALFAFMSMALELOGAIISWEAPVEKTECIQKNNQTCFNFIRLOPY 173
OY 121 NASHLYVCGTVAFOPKCTVYNNMLFTLEHGEFEDGKGCPRYPANAGHLLVDGELYSAT 180
DB 121 NASHLYVCGTVAFOPKCTVYNNMLFTLEHGEFEDGKGCPRYPANAGHLLVDGELYSAT 180
OY 174 NASHLYVCGTVAFOPKCTVYNNMLFTLEHGEFEDGKGCPRYPANAGHLLVDGELYSAT 233
DB 174 NASHLYVCGTVAFOPKCTVYNNMLFTLEHGEFEDGKGCPRYPANAGHLLVDGELYSAT 233
OY 181 LNPFLETEPIILRNKGPBHSMTETYLAFWLNPHFVGSAYVESVGSFSGDDDKYFFER 240
DB 181 LNPFLETEPIILRNKGPBHSMTETYLAFWLNPHFVGSAYVESVGSFSGDDDKYFFER 240
OY 234 LNPFLETEPIILRNKGPBHSMTETYLAFWLNPHFVGSAYVESVGSFSGDDDKYFFER 293
DB 234 LNPFLETEPIILRNKGPBHSMTETYLAFWLNPHFVGSAYVESVGSFSGDDDKYFFER 293
OY 241 ERAVESDCYAEQVAVARAVARVCKDGMGARTLQKMTTFELKARLACAPMMQLYFNLOLM 300
DB 241 ERAVESDCYAEQVAVARAVARVCKDGMGARTLQKMTTFELKARLACAPMMQLYFNLOLM 300
OY 294 ERAVESDCYAEQVAVARAVARVCKDGMGARTLQKMTTFELKARLACAPMMQLYFNLOLM 353
DB 294 ERAVESDCYAEQVAVARAVARVCKDGMGARTLQKMTTFELKARLACAPMMQLYFNLOLM 353
OY 301 HTLODTSMHNTTFFGVFOAQMGDMYLSAICEYOLEEIQVFEGPYKEVHEEAQKMDRYD 360
DB 301 HTLODTSMHNTTFFGVFOAQMGDMYLSAICEYOLEEIQVFEGPYKEVHEEAQKMDRYD 360
OY 354 HTLODTSMHNTTFFGVFOAQMGDMYLSAICEYOLEEIQVFEGPYKEVHEEAQKMDRYD 413
DB 354 HTLODTSMHNTTFFGVFOAQMGDMYLSAICEYOLEEIQVFEGPYKEVHEEAQKMDRYD 413

```

QY 361 PVSPRPGSCINNNRRHGGTSSLELDPDNIINFTYKKHPLMEEOYGPMSRPLLYKKGTNF 420
DB 414 PVSPRPGSCINNNRRHGGTSSLELDPDNIINFTYKKHPLMEEOYGPMSRPLLYKKGTNF 473
QY 421 THVADRVNTGIDGATYVLTFTIGDGMILKAVSLGPPWHLLLEQLPFOEPMRSLVLSQS 480
DB 474 THVADRVNTGIDGATYVLTFTIGDGMILKAVSLGPPWHLLLEQLPFOEPMRSLVLSQS 533
QY 481 KKLIFAGRSQSLVOLPVADCIKYRSCADCVIARDPYCAMSVNTSRCAVAGHSGSLIIQH 540
DB 534 KKLIFAGRSQSLVOLPVADCIKYRSCADCVIARDPYCAMSVNTSRCAVAGHSGSLIIQH 593
QY 541 VMTSDTSGICNLKRGSKTQSGPXPKNITVYAGTDVLVPCHLSSNLA-----LPDSN 590
DB 594 VMTSDTSGICNLKRGSK-KVREPTPKNITVYAGTDVLVPCHLSSNLAHAMWTFGRDLPAEQ 652
QY 591 P 591
DB 653 P 653

RESULT 10
AAB61238
ID AAB61238 standard; Protein: 834 AA.
XX
AC AAB61238;
XX
DT 03-APR-2001 (first entry)
XX
DE Murine M-Sema-F protein.
XX
KW Mouse; M-Sema-F; INTERCEPT 217; INTERCEPT 297; TANGO 276;
KW TANGO 292; TANGO 325; TANGO 331; TANGO 332; cytoskeletal; antiinflammatory;
KW antiarrhythmic; antipsoriatic; gene therapy; cancer;
KW inflammatory disorder; cardiac disorder; arrhythmia; skin disorder;
KW psoriasis.
XX
OS Mus sp.
XX
PN WO200100638-A2.
XX
PD 04-JAN-2001.
XX
PF 16-JUN-2000; 2000MO-US16658.
XX
PR 29-JUN-1999; 99US-0342364.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Kitist SJ, Holtzman DA, Fraser CC, Sharp JD, Barnes TS;
XX
DR WPI: 2001-061966/07.
XX
N-PSDB; AAF29461.
XX
PT Isolated human proteins are used for diagnosis, treatment and
PT prevention of cancers, inflammatory disorders, cardiac disorders e.g.
PT arrhythmia, and skin disorders e.g. psoriasis.
XX
PS Disclosure: Fig 3F-3H; 372pp; English.
XX
XX
CC The present sequence is given in a specification relating to isolated
CC human proteins designated INTERCEPT 217, INTERCEPT 297, TANGO 276,
CC TANGO 292, TANGO 325, TANGO 331 and TANGO 332. These proteins are useful
CC as modulating agents or as targeting agents for developing agents to
CC regulate cellular processes e.g. growth, proliferation, survival,
CC differentiation and activity of human tissues. Diseases which can be
CC diagnosed, prevented and treated by administration of these polypeptides,
CC their nucleic acids and modulators include cancers, inflammatory
CC disorders, cardiac disorders e.g. arrhythmia, and skin disorders e.g.
CC psoriasis. Nucleic acids encoding the isolated proteins can be used to
CC express the proteins in a host cell in gene therapy applications.
CC Antisense molecules or ribozymes can be used to inhibit expression
CC of the proteins in target cells. Fragments of the nucleic acid molecules

```

```

CC encoding the isolated proteins are used as hybridisation probes and as
CC polymerase chain reaction (PCR) primers.
XX
SQ Sequence 834 AA;
QY 1 MAPHAAVLLAARLMGICIGAEWMNNVPRKTVSSGELATVRRPSQIDDFLTITTE 60
DB 1 MAPHAAVLLAAGLMLGIGAEWMNNVPRKTVSSGELATVRRPSQIDDFLTITTE 60
QY 61 PTGLIYVAREBALFAFSHEALELQGAISMEAPVEKTEICIOGKNNOTECFNFIFLOPY 120
DB 61 HSGILYVAREBALFAFSHEALELQGAISMEAPVEKTEICIOGKNNOTECFNFIFLOPY 120
QY 121 NASHLYVCGTYAFOPKCTYVNMILFTLEHGEFEDGKGCOPYDPAKGAAGLVDELYSAT 180
DB 121 NASHLYVCGTYAFOPKCTYVNMILFTLEHGEFEDGKGCOPYDPAKGAAGLVDELYSAT 180
QY 181 LNPLFTEPITLRNNGPHHSMKTEYLAFMLNPHVGSAAVPEVSGFTGDDKRYEPR 240
DB 181 LNPLFTEPITLRNNGPHHSMKTEYLAFMLNPHVGSAAVPEVSGFTGDDKRYEPR 240
QY 241 ERAVEDCYADQVAVARAVCKGDMGARTLQKWTTFELKARLACSAFNMOLYFNQLOAM 300
DB 241 ERAVEDCYADQVAVARAVCKGDMGARTLQKWTTFELKARLACSAFNMOLYFNQLOAM 300
QY 301 HTLQDTSMHNTTFEGVFQAOQWMDTLAICEYQLEIQRVFEQPYKEYHEEAQKMDRYTD 360
DB 301 HTLQDTSMHNTTFEGVFQAOQWMDTLAICEYQLEIQRVFEQPYKEYHEEAQKMDRYTD 360
QY 361 PVSPRPGSCINNNRRHGGTSSLELDPDNIINFTYKKHPLMEEOYGPMSRPLLYKKGTNF 420
DB 361 PVSPRPGSCINNNRRHGGTSSLELDPDNIINFTYKKHPLMEEOYGPMSRPLLYKKGTNF 420
QY 421 THVADRVNTGIDGATYVLTFTIGDGMILKAVSLGPPWHLLLEQLPFOEPMRSLVLSQS 480
DB 421 THVADRVNTGIDGATYVLTFTIGDGMILKAVSLGPPWHLLLEQLPFOEPMRSLVLSQS 480
QY 481 KKLIFAGRSQSLVOLPVADCIKYRSCADCVIARDPYCAMSVNTSRCAVAGHSGSLIIQH 539
DB 481 KKLIFAGRSQSLVOLPVADCIKYRSCADCVIARDPYCAMSVNTSRCAVAGHSGSLIIQH 540
QY 540 HMTSDTSGICNLRG-SKIOSGPXPKNITVYAGTDVLVPCHLSSNLA-----LPD 588
DB 540 HMTSDTSGICNLRG-SKIOSGPXPKNITVYAGTDVLVPCHLSSNLAHAMWTFGRDLPA 598
QY 589 SNP 591
DB 599 EQP 601

RESULT 11
ABU60951
ID ABU60951 standard; Protein: 785 AA.
XX
AC ABU60951;
XX
DT 08-MAY-2003 (first entry)
XX
DE Lung specific protein (LSP) #54.
XX
KW Human; gene therapy; vaccine; lung specific antigen; cancer diagnosis;
KW cancer monitoring; cancer staging; cancer imaging; lung cancer;
KW non-cancerous diseases of the lung; transgenic animal.
XX
OS Homo sapiens.
XX
PN WO2002068633-A2.
XX
PD 06-SEP-2002.

```


Db 244 YFNOLAMHTLQDTSMHNTTFEGVFOAQKGDMLSAICEYOLEIQRVEGPIKEHEA 303
 QY 353 OKMDRYTDPVSPSPGSCINNMHRRHGYTSSLELPDNLNFVKKHPIMEQVGRPMSRPL 412
 Db 304 OKMDRYTDPVSPSPGSCINNMHRRHGYTSSLELPDNLNFVKKHPIMEQVGRPMSRPL 363
 QY 413 LVKKGTFHTLVADRYTGDLGATYTVLFITGDMILKAVSLGFWHLIEQLDFDEPM 472
 Db 364 LVKKGTFHTLVADRYTGDLGATYTVLFITGDMILKAVSLGFWHLIEQLDFDEPM 423
 QY 473 RSLVLSQSK-KLIFAGRSQVLVOLPVADCIKYSRSCACVILADPYCAMSVNTSRCAVVG 531
 Db 424 RSLVLSQSKKLLPAGRSQVLVOLPVADCIKYSRSCACVILADPYCAMSVNTSRCAVVG 483
 QY 532 HFGSLLIQHWMTSDTSGICNLRGSKIOSGPKNITVAGTDLVPCHLSSNLA----- 585
 Db 484 HSGSLLIQHWMTSDTSGICNLRGSK-KVRPTPKNITVAGTDLVPCHLSSNLAHAWTF 542
 QY 586 ----LPPSNP 591
 Db 543 GGRDLPAEQP 552

RESULT 13

ABB97213 ID ABB97213 standard; Protein: 569 AA.

XX ABB97213;

DT 28-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 481.

XX Human; antihaemic; vulnery; antiinflammatory; immunomodulator;
 KW antifertility; cerebroprotective; cyrostatic; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; EST;
 KW expressed sequence tag.

XX Homo sapiens.

XX WO200222660-A2.

XX 21-MAR-2002.

PF 10-SEP-2001; 2001WO-US26015.

PR 11-SEP-2000; 2000US-0659671.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhang QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

DR WPI: 2002-292408/33.

DR N-PSDB; ABN32399.

PT An isolated polynucleotide for treating diseases associated with its

XX encoded polypeptide such as cancer and multiple sclerosis -

PS Example 2; SEQ ID NO 481; 509pp; English.

XX The present invention provides the protein and coding sequences of 444

CC novel human proteins. These were isolated from expressed sequences tags

CC (ESTs). They can be used to stimulate cell growth, to regulate

CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth

CC e.g. in burn treatment, to regulate the immune system e.g. to treat

CC multiple sclerosis, to regulate activin or inhibin e.g. to treat

CC infertility, to regulate haemostasis or thrombolysis e.g. to treat

CC stroke and cancer, to screen for drugs, to treat inflammatory conditions

CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.

CC Parkinson's disease. The present sequence is a protein of the invention.

XX Sequence 569 AA;

Query Match 52.4%; Score 1685.5; DB 23; Length 569;
 Best Local Similarity 93.8%; Pred. No. 1.3e-161;
 Matches 316; Conservative 2; Mismatches 8; Indels 11; Gaps 2;

QY 265 MGAQRITQKRWTTFLKARLACSAFNWOLYFNOLAMHTLQDTSMHNTTFEGVFOAQKGD 324
 Db 1 MGAQRITQKRWTTFLKARLACSAFNWOLYFNOLAMHTLQDTSMHNTTFEGVFOAQKGD 60
 QY 325 YLSAICEYOLEIQRVEGPIKEHEAOKMDRYTDPVSPSPGSCINNMHRRHGYTSS 384
 Db 61 YLSAICEYOLEIQRVEGPIKEHEAOKMDRYTDPVSPSPGSCINNMHRRHGYTSS 120
 QY 385 ELPDNLNFVKKHPIMEQVGRPMSRPLVYKKGTFHTLVADRYTGDLGATYTVLFITG 444
 Db 121 ELPDNLNFVKKHPIMEQVGRPMSRPLVYKKGTFHTLVADRYTGDLGATYTVLFITG 180
 QY 445 DGRLLKAVSLGFWHLIEQLDFDEPMRSLSVLSQSKKLLPAGRSQVLVOLPVADCIK 504
 Db 181 DGRLLKAVSLGFWHLIEQLDFDEPMRSLSVLSQSKKLLPAGRSQVLVOLPVADCIK 240
 QY 505 SCADCVILADPYCAMSVNTSRCAVVGHFGSLLIQHWMTSDTSGICNLRGSKIOSGPK 564
 Db 241 SCADCVILADPYCAMSVNTSRCAVVGHFGSLLIQHWMTSDTSGICNLRGSK-KVRPTPK 299
 QY 565 NITVAGTDLVPCHLSSNLA-----LPPSNP 591
 Db 300 NITVAGTDLVPCHLSSNLAHAWTFGGRDLPAEQP 336

RESULT 14

AAE03818 ID AAE03818 standard; Protein: 832 AA.

XX AAE03818;

DT 08-AUG-2001 (first entry)

DE Human gene 1 encoded secreted protein HRAHL26, SEQ ID NO: 64.

KW Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
 KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
 KW inflammation; neurological disorder; Alzheimer's disease; food additive;
 KW angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
 KW pregnancy-related disorder; endocrine disorder; infection; wound healing;
 KW cell culture; chemotaxis; vulnery; binding partner identification;
 KW gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 1..38

XX Protein 39..832

XX /note="Mature secreted protein"

XX WO200136440-A1.

XX 25-MAY-2001.

XX 15-NOV-2000; 2000WO-US31282.

XX 19-NOV-1999; 99US-0166414.

XX 21-JUL-2000; 2000US-0219665.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Komatsoulis GA, Birse CE, Moore PA;

XX WPI: 2001-343795/36.

XX

DR

DR N-PSDB: AAD08283.

Isolated nucleic acid molecule encoding a human secreted protein is used in preventing, treating or ameliorating a medical condition

Claim 11, Page 476-479; 553pp; English.

CC AAD08283-AAD0335 represent cDNAs corresponding to 23 human secreted
CC protein genes, and AAE03818-AAE03870 represent the proteins they encode.
CC AAE03871-AAE03896 represent human secreted protein fragments or variants.
CC The secreted proteins and their genes are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 23 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumours, foetal and developmental
CC abnormalities, hematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angiogenic disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunoassay (ELISA). The present sequence represents a human
CC secreted protein of the invention.

SO Sequence 832 AA:

Query Match 38.5%; Score 1240; DB 22; Length 832;

Best Local Similarity 45.4%; Pred. No. 4,3e-116;
Matches 265; Conservative 88; Mismatches 201; Indels 30; Gaps 12;

QY 25 WLVVRKTYSSGLATVVRFRFSGTIGDPLTITLTPRGLLVGAREALFAFSMALEIQ 84
DB 37 WALSPRISPLGSEERRPLRFAEHI NNTALLSRGRTLVYGAREALFAFSMALEIQ 96
QY 85 GA-----ISWEAFVEKTECTIOGKNNQTECFNIFRLOPYNASHLYVCTVAFOPRCTYV 140
DB 97 GGEYDELMGADAEKKQCCSFGRKDPORCQYIILPLSGSHLFTGTAAFSMCTYI 156
QY 141 NMLTFTLEHGE-----FEDGKCKCPYDPAKGHAGLLVDGELYSATLNFLCTEPTILRM 195
DB 157 NNEETFLARDEKGNVLEDKGRCPCFDPNFKSTALVYDGEELTVGVSSFGQDPAISRQ 216
QY 196 GPHNMTKEEYLAFLNNEHVFVSAYVPSVSGFTGDDKVFYFFERAVESCIYEQVYA 255
DB 217 SLRPT-KTESSLNWLODPAFVASAVIPESLSLODDBDIYFFESTQCEFEFEFTYIS 275
QY 256 RYAVRCKDMDGARTLQKMTFTLKARLACAPNOLVFNOAHHTLODT--SMHNTVF 313
DB 276 RIARCKDDEGERVLOQRWTSFLNQLLCSRFDDGFPFNVDYFTLSPSQDMRDLTF 335
QY 314 FGVFOAQN--GDWYLSAICEVQLEIQRVEGPEYEHEDAKMDRYDTPVSPRPGSCI 371
DB 336 YGVFTSOMHRTGTEBSAVCFMVKDYQRFVSLYKEVRETQOQWTVTHPVPTRPGACI 395
QY 372 NMMHRRIGVTSLELPDNLINLVKHPMLMEBYGGRMSPLLYKKGTFNTHLYADRVGL 431
DB 396 TNSARERRINSSLOLPDRVLFNFKDHFMDGQVR--SMMLLQDQARQORAVAVHVRPL 452
QY 432 DGATVYLFIGTGDMLKAVSLGPNVHLIELQLEFDD-EPARSVLVSQSKTLFAGSRS 490
DB 432 DGATVYLFIGTGDMLKAVSLGPNVHLIELQLEFDD-EPARSVLVSQSKTLFAGSRS 490

DB 453 H-HTVDVFLGTCGDLHLKRAVSGVPRVHIIIEELQIFSSGQPVONLLDTHRGILLVYASHS 511
QY 491 QLVLPVADCIKRYSCADCLADPYCAMSVMNTRCAVAGHPSGLL----IQHWMTSDT 546
DB 512 GVVQVPMANCSLRSRGDCDLADPYCAMS--GSSCKHVSIVPOLATRWIDIDIGASA 569
QY 547 SGICN-----LRGSKIOSGPPR-KNITVVAAGTDLVLPCHLSSNIA 585
DB 570 KDLCSASVSPSPFVPTGKPCBOVOPFPNTVNTLACPLLSNIA 613

RESULT 15

ID ABG64522 standard; Protein; 832 AA.

AC ABG64522;

DT 27-ANG-2002 (first entry)

DE Human albumin fusion protein #1197.

KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytostatic; antifertility; antiinflammatory; anticancer;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neurotropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.

OS Homo sapiens.

OS Synthetic.

PN WO20017137-A1.

PD 18-OCT-2001.

PF 12-APR-2001; 2001WO-US11988.

PR 12-APR-2000; 2000US-229358P.

PR 25-APR-2000; 2000US-199384P.

PR 21-DEC-2000; 2000US-256931P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Haseltine WA;

PI WPI; 2002-010886/01.

XX New fusion protein for treating disease e.g. diabetes comprises an

PT albumin fused to a therapeutic protein -

PS Claim 1; Page 1288-1291; 2102pp; English.

XX The present invention relates to albumin fusion proteins comprising a

CC therapeutic protein X and human albumin (HA, also known as human serum

CC albumin, HSA). The proteins are useful for treating a disease or

CC disorder that may be modulated by therapeutic protein X. The albumin

CC extends the shelf-life of protein X, and may increase its biological

CC in vitro/in vivo activity. The protein is useful for treating and

CC diagnosing disorders such as cancer, reproductive disorders, digestive

CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders

CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders

CC (e.g. diabetes), haematopoietic disorders, neural disorders

CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,

CC encephalomyelitis, meningitis, schizophrenia), and connective disorders

CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin

fusion proteins of the invention.

SO Sequence 832 AA:
Query Match 38.5%; Score 1240; DB 23; Length 832;
Best Local Similarity 45.4%; Pred. No. 4,3e-116;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2003, 13:42:14 ; Search time 26 Seconds
(without alignments)
1077.997 Million cell updates/sec

Title: US-10-002-050-14

Perfect score: 3219
Sequence: 1 MAPHAWVLLAARLWGLGIG.....PCHLSSNALPDSPNESSV 596

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2720	84.5	SM4C_MOUSE	064151 mus musculus
2	1240	38.5	SM4B_HUMAN	09np22 homo sapien
3	1225.5	38.1	SM4G_HUMAN	09nln3 homo sapien
4	1223.5	38.0	SM4G_MOUSE	09nln3 mus musculus
5	1222.5	38.0	SM4D_MOUSE	09nln3 mus musculus
6	1219.5	37.9	SM4D_HUMAN	092854 homo sapien
7	1149	35.7	SM4B_MOUSE	062179 mus musculus
8	1146.5	35.6	SM27_BRARE	09nln3 brachydanto
9	1029.5	32.0	SM4A_HUMAN	09nln3 homo sapien
10	1008.5	31.3	SM4F_RAT	092143 rattus norv
11	1005.5	31.2	SM4F_HUMAN	095754 homo sapien
12	1000.5	31.1	SM4F_MOUSE	092123 mus musculus
13	992.5	30.8	SM4A_MOUSE	062178 mus musculus
14	937.5	29.1	SM3A_MOUSE	008665 mus musculus
15	937.5	29.1	SM3A_RAT	093548 rattus norv
16	934.5	29.0	SM3A_CHICK	090607 gallus gall
17	932.5	29.0	SM3A_HUMAN	014563 homo sapien
18	913	28.4	SM2B_BRARE	094686 brachydanto
19	902.5	28.0	SM3D_CHICK	090663 gallus gall
20	895.5	27.8	SM3D_HUMAN	095025 homo sapien
21	892	27.7	SM3B_HUMAN	013214 homo sapien
22	890	27.6	SM2A_BRARE	094731 brachydanto
23	872.5	27.1	SM3B_MOUSE	062177 mus musculus
24	844.5	26.2	SM3E_HUMAN	015041 homo sapien
25	841.5	26.1	SM3E_MOUSE	010275 mus musculus
26	837	26.0	SM22_BRARE	094696 brachydanto
27	828	25.7	SM3F_MOUSE	086632 mus musculus
28	827	25.7	SM3F_HUMAN	013275 homo sapien
29	804.5	25.0	SM3E_CHICK	042237 gallus gall
30	793.5	24.7	SM3C_CHICK	042236 gallus gall
31	779	24.2	SM3C_HUMAN	090985 homo sapien
32	774	23.0	SM3C_MOUSE	062181 mus musculus
33	739	23.0	SM1A_CAHEL	017330 caenorhabdi

34	730	22.7	771	1	SM1A_DROME	024322 drosophila
35	729.5	22.7	1074	1	SM5A_HUMAN	013591 homo sapien
36	729	22.6	1077	1	SM5A_MOUSE	062217 mus musculus
37	726.5	22.6	930	1	SM6C_HUMAN	09h312 homo sapien
38	719.5	22.4	295	1	SM4D_CHICK	090665 gallus gall
39	708.5	22.0	730	1	SM1A_CHICK	026473 schistocerc
40	708	22.0	1093	1	SM5B_MOUSE	060519 mus musculus
41	706.5	21.9	931	1	SM6C_MOUSE	09wt13 mus musculus
42	704.5	21.9	960	1	SM6C_RAT	09wt13 rattus norv
43	704	21.9	697	1	SM2A_CHICK	09x268 schistocerc
44	686	21.3	1030	1	SM6A_HUMAN	09h266 homo sapien
45	674	20.9	888	1	SM6A_MOUSE	035464 mus musculus

ALIGNMENTS

RESULT 1
SM4C_MOUSE STANDARD: PRT: 834 AA.
ID SM4C_MOUSE
AC 064151;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 4C precursor (Semaphorin I) (Sema I) (Semaphorin C-like 1)
DE (M-Sema F).
GN SEMA4C OR SEMA1 OR SEMA1L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Neonatal brain;
RX MEDLINE=95385809; PubMed=7656991;
RA Inagaki S., Furuyama T., Iwashita Y.;
RT Identification of a member of mouse semaphorin family.*;
RL FEBS Lett. 370:269-272(1995).
RN [2]
RP INTERACTION WITH GIPC.
RX MEDLINE=99253973; PubMed=10318831;
RA Wang L.-H., Kald R.G., Stittmayer S.M.;
RT "A PDZ protein regulates the distribution of the transmembrane
semaphorin, M-SemF".
RL J. Biol. Chem. 274:14137-14146(1999).
CC -1- SUBUNIT: INTERACTS WITH GIPC PDZ DOMAIN.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED WIDELY IN THE NERVOUS TISSUES DURING
DEVELOPMENT. PREDOMINANTLY IN THE DEVELOPING BRAIN AND SPINAL
CORD.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: Contains 1 Sema domain.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC EMBL: S79463; AAB35184.1; -
DR PIR: S66498; S66498.
DR MCD: MGI:109252; Sema4C.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR002165; Plexin-repeat.
DR InterPro: IPR001627; Sema.
DR Pfam: PF01437; PSI; 1.
DR Pfam: PF01403; Sema; 1.

DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS00835; IG-LIKE; 1.
 KM Signal; Transmembrane; Immunoglobulin domain; Multigene family;
 KM Neurogenesis; Developmental protein; Glycoprotein.
 FT SIGNAL 1 20
 FT CHAIN 1 20
 FT DOMAIN 21 834
 FT TRANSMEM 21 664
 FT DOMAIN 665 685
 FT DOMAIN 686 834
 FT DOMAIN 53 481
 FT DOMAIN 489 552
 FT DOMAIN 557 645
 FT DOMAIN 678 687
 FT DOMAIN 709 774
 FT DOMAIN 831 834
 FT DISULFID 578 628
 FT CARBOHYD 106 106
 FT CARBOHYD 121 121
 FT CARBOHYD 310 310
 FT CARBOHYD 419 419
 FT CARBOHYD 522 522
 FT CARBOHYD 565 565
 SQ SEQUENCE 834 AA; 92556 MW; 6668B5B5F71482D CRC64;

Query Match 84.5%; Score 2720; DB 1; Length 834;
 Best Local Similarity 83.1%; Pred. No. 6,4e-22;
 Matches 501; Conservative 41; Mismatches 47; Indels 14; Gaps 4;

1 MAPHAWMLLAARGLGIGAEVMMNIVPRKTVSSGELATVVRSSQGIQDFITLTLE 60
 1 MAPHAWMLLAARGLGIGAEVMMNIVPRKTVSSGELATVVRSSQGIQDFITLTLE 60
 61 PTGLLYVAREALFAFSEMELEQAIISWEAVEKEKTECIQKGNQTECFNFIETLOPY 120
 61 HSGLLYVAREALFAFSEMELEQAIISWEAVEKEKTECIQKGNQTECFNFIETLOPY 120
 121 NASHLYVGTAFVFOKCTYVNNLTFTLEGEFEDGKCKCPYPANAGHGLVYDGLYSAT 180
 121 NASHLYVGTAFVFOKCTYVNNLTFTLEGEFEDGKCKCPYPANAGHGLVYDGLYSAT 180
 121 NSSHLYVGTAFVFOKCTYVNNLTFTLEGEFEDGKCKCPYPANAGHGLVYDGLYSAT 180
 181 LNPFGLTEPIILRNKGPVHSMTKTEFLAFVLPNPFVGSATYVESYGFDDDKYFEFR 240
 181 LNPFGLTEPIILRNKGPVHSMTKTEFLAFVLPNPFVGSATYVESYGFDDDKYFEFR 240
 181 LNPFGLTEPIILRNKGPVHSMTKTEFLAFVLPNPFVGSATYVESYGFDDDKYFEFR 240
 181 LNPFGLTEPIILRNKGPVHSMTKTEFLAFVLPNPFVGSATYVESYGFDDDKYFEFR 240
 241 ERAVSDCYAEQVAVARAVKCGDKGARTLOKRTTFLKARLACAPMQLYFNOLOQM 300
 241 ERAVSDCYAEQVAVARAVKCGDKGARTLOKRTTFLKARLACAPMQLYFNOLOQM 300
 241 ERAVSDCYAEQVAVARAVKCGDKGARTLOKRTTFLKARLACAPMQLYFNOLOQM 300
 241 ERAVSDCYAEQVAVARAVKCGDKGARTLOKRTTFLKARLACAPMQLYFNOLOQM 300
 301 HTLRASVSHNMTTFEFCVFOARWGDMDISAVCEYOLIOOVFEGPKREYSECAOKARAYTD 360
 301 HTLRASVSHNMTTFEFCVFOARWGDMDISAVCEYOLIOOVFEGPKREYSECAOKARAYTD 360
 361 PVPSPRPGSCINNMHRHGYTSLELPNINLVFKKHPLEMEQVCPKMSRPLLYKKNF 420
 361 PVPSPRPGSCINNMHRHGYTSLELPNINLVFKKHPLEMEQVCPKMSRPLLYKKNF 420
 361 PVPSPRPGSCINNMHRHGYTSLELPNINLVFKKHPLEMEQVCPKMSRPLLYKKNF 420
 361 PVPSPRPGSCINNMHRHGYTSLELPNINLVFKKHPLEMEQVCPKMSRPLLYKKNF 420
 421 THLVADRYTGLDGAATYTLFICTGDMILKAVSLGSPVHLIELDLFDEPARSLVLSOS 480
 421 THLVADRYTGLDGAATYTLFICTGDMILKAVSLGSPVHLIELDLFDEPARSLVLSOS 480
 421 THLVADRYTGLDGAATYTLFICTGDMILKAVSLGSPVHLIELDLFDEPARSLVLSOS 480
 421 THLVADRYTGLDGAATYTLFICTGDMILKAVSLGSPVHLIELDLFDEPARSLVLSOS 480
 481 KKLVLGAGRSOLVLPVADCIYRSCADCVIARDPYCAWSVNTSCVA-VGSHFSLIQ 539
 481 KKLVLGAGRSOLVLPVADCIYRSCADCVIARDPYCAWSVNTSCVA-VGSHFSLIQ 539
 481 KKLVLGAGRSOLVLPVADCIYRSCADCVIARDPYCAWSVNTSCVA-VGSHFSLIQ 539
 481 KKLVLGAGRSOLVLPVADCIYRSCADCVIARDPYCAWSVNTSCVA-VGSHFSLIQ 539
 540 HMTSDTSGICNLRG-SKIOSGPKRNTYVAGTDLVPCCHSSNLA-----LPD 588
 540 HMTSDTSGICNLRG-SKIOSGPKRNTYVAGTDLVPCCHSSNLA-----LPD 588
 541 HVAHNDTSMCNQIYIKVRS--IPKNITVSGTDLVPCCHSSNLAHVAHMTFGSODLPA 598
 589 SNP 591
 599 BOP 601

RESULT 2
 ID SM4B_HUMAN STANDARD; PRT; 832 AA.
 AC O9NPR2; O9NPR9; O9NPR5; O9NPR8; O9NPR1; O9NPR6; O9NPR0;
 DT 16-OCT-2001 (rel. 40, Created)
 DT 15-SEP-2003 (rel. 42, Last sequence update)
 DT 15-SEP-2003 (rel. 42, Last annotation update)
 DE Semaphorin 4B precursor.
 GN SEM4B OR KIAA1745.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=21082932; PubMed=11214970;
 RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
 RT The complete sequences of the coding sequences of unidentified human genes. XIX.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 7:347-355(2000).
 RN [2]
 RP SEQUENCE OF 160-832 FROM N.A., AND VARIANT ALA-792.
 RA Carim L., Estivill X., Escarceller M., Sumoy L.;
 RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 358-832 FROM N.A.
 RA Kawabata A., Hiki T., Kobatake N., Inagaki H., Iyema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Okeyashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Isegai T., Sugano S.;
 RT "NEO human cDNA sequencing project."
 RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 303-832 FROM N.A., AND VARIANT ALA-792.
 RC TISSUE=Colon, and Pancreas;
 RX MEDLINE=22389257; PubMed=12477932;
 RA Strusberg R.L., Felngold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Hsieh F.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Matovina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loguelfano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield J.S.N., Krzywicki M.T., Skalska U., Smallus D.E.,
 RA Schermer A., Schein J.E., Jones S.J.M., Maitra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
 CC SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: Contains 1 Sema domain.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -1- CAUTION: Ref.3 sequence differs from that shown in position 709
 CC onward due to a frameshift.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AB051532; BAB21836.1; -
 DR EMBL; AL390080; CAB98204.1; -
 DR EMBL; AL390081; CAB98205.1; -
 DR EMBL; AL390082; CAB98206.1; -
 DR EMBL; AK026133; BAB15372.1; ALT_FRAME.
 DR EMBL; BC010701; AAH10701.1; ALT_INIT.
 DR EMBL; BC01658; AAH17658.1; -
 DR Genew; HGNC:10730; SEMA4B.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR002165; Plexin_repeat.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01437; PSI; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 KW Transmembrane; Immunoglobulin domain; Multigene family; Neurogenesis;
 KM Developmental; protein; glycoprotein; signal; polymorphism.
 FT SIGNAL 1 38
 FT CHAIN 39 832
 FT DOMAIN 39 712
 FT TRANSMEM 713 733
 FT DOMAIN 734 832
 FT DOMAIN 65 502
 FT DOMAIN 520 574
 FT DOMAIN 599 658
 FT DOMAIN 753 776
 FT DISULFID 606 651
 FT CARBOHYD 64 64
 FT CARBOHYD 91 91
 FT CARBOHYD 160 160
 FT CARBOHYD 405 405
 FT CARBOHYD 520 520
 FT CARBOHYD 625 625
 FT VARIANT 792 792
 FT CONFLICT 160 160
 FT CONFLICT 565 565
 SQ SEQUENCE 832 AA; 92192 MW; 29D58C1DD5E1C6B CRC64;

Query Match 38.5% Score 1240; DB 1; Length 832;
 Best Local Similarity 45.4%; Pred. No. 2.6e-96;
 Matches 265; Conservative 88; Mismatches 201; Indels 30; Gaps 12;

QY 25 WNLVPRVYSSGELATVVRPSQIGDDELTLITETGLLYVCAKALPAFSEALEIQ 84
 DB 37 WALSPRISLPGLSEERPLRFEEAHISNYTALISRDGRITLYGAREALFALSINLFLP 96
 QY 85 GA-----ISWEAPVAKTECIQKGNKNOTECFNFRILOPYNAVSLYVCGYAFQPKCTYV 140
 DB 97 GGEVQELLMGDADEKAKQCKSKGNDPORDCONYIKILPLSGSHLFTCGTAASFPCYTI 156
 QY 141 NMLTFLEHGE-----FEDGKGCPRYPDAGHAGLVNDELVSATLNFLTEBETILNM 195
 DB 157 NMENFTLARDKGNVLEDEKGRCPDPNFKSTALVVDGLTYGVTSFQGNDAISRSQ 216
 QY 196 GPHHSMTETYLAFYLANPHEVGSAYVPSVGSFTGDDDKYFFERRAVERSDCAEOVYA 255
 DB 217 SLRPT-FTESSLNLDPAFVASAYIPESLSLOGDDDKYFFESSEGOEFEPENITVS 275
 QY 256 RVAVSCGKDGAGATTLQKRTTFPKARLACAPRNQLOFNLOLQAMHTLODT--SMHNTF 313
 DB 276 RIRACIGDEGEGEVNLDKRTSFLKADLLCSRPDGGFPFNVLQVFTLSPEPQMRDLTF 335
 QY 314 FGVEQAQW--GDMTLSAICEVLEEIQVTEGPRKEYHEAQAQKDRYTPVPSPSCSI 371
 DB 336 YGVTSOMHNGTTEGSAVCFYTKMDVOVRFSGLYKVNRETOQVYTVTHVPPTPRPACI 395
 QY 372 NNHRRRGGTSSLEPNNILNFAVFKHPLMEQVQPRKSRPLLYVKGKGNFTHLVADRVTGL 431

DB 396 TNSAREKINSIQLPDVNLFLKDHFLMDQVR---SRMLLLQPOARYORAVHVRPGL 452
 QY 432 DGATYVLEIGTGGMILKAVSLGPRVHLIEQLDFD--EPMRSLVTSQSKLLPAGRS 490
 DB 453 H-HYVDVLEFGTGGRKLKAVSVGRVHIIIELOIFSSGOVYVNLIDTNGHLTPASHS 511
 QY 491 QLVDLPVADCIKYSACDVLARDPYCAMSVNTRCAVAGHFGSLT-----TQHVMTSDT 546
 DB 512 GVVQVPMVNCISLYSCGCLLARDPYCAMS--GSSCKIVSLYQPLATRPIDIDEGASA 569
 QY 547 SGICN-----LNGSKIQSGPR-KNITYVAGTDVLPCLSSNLA 585
 DB 570 KDLCSASSVSPSPVPTGKPCQVQVOPQNTVNTLACPLLSNLA 613

RESULT 3
 SM4G_HUMAN STANDARD; PRT; 838 AA.
 ID SM4G_HUMAN
 AC O9NTN9; O9HCF3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 4G precursor.
 GN SEMA4G OR KIA1619.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Heath P.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20450683; PubMed=1097877;
 RA Nagase T., Kikuno R., Nakayama M., Hirotsawa M., Ohara O.;
 RT *Prediction of the coding sequences of unidentified human genes.
 RT XVIII. The complete sequences of 100 new cDNA clones from brain which
 RT code for large proteins in vitro.*;
 RL DNA Res. 7:273-281(2000).
 CC CC
 CC -1- FUNCTION: MAY PLAY A ROLE IN AXON GUIDANCE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: Contains 1 Sema domain.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AL133215; -; NOT_ANNOTATED_CDS.
 CC EMBL; AB046839; BAB13445.1; ALT_INIT.
 CC Genew; HGNC:10735; SEMA4G.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003599; Ig.
 CC InterPro; IPR003006; Ig_MHC.
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR002165; Plexin_repeat.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF01437; PSI; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00409; IG; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC PROSITE; PS0835; IG-like; 1.
 KW Signal; Transmembrane; Immunoglobulin domain; Multigene family;
 KM Neurogenesis; Developmental protein; glycoprotein.
 FT SIGNAL 1 17
 FT CHAIN 18 838
 FT SEMAPHORIN 4G.

FT DOMAIN 18 675 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 676 686 POTENTIAL.
 FT DOMAIN 697 838 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 56 489 SEMA.
 FT DOMAIN 507 558 PSI.
 FT DOMAIN 567 649 IG-LIKE C2-TYPE.
 FT DOMAIN 565 568 POLY-PRO.
 FT DOMAIN 763 774 POLY-PRO.
 FT DISULFID 584 632 BY SIMILARITY.
 FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 598 598 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 543 543 N-> RSQGR (IN REF. 2).
 SQ SEQUENCE 838 AA; 91496 MW; 98281AEE861F245 CRC64;

Query Match 38.1%; Score 1225.5; DB 1; Length 838;
 Best Local Similarity 46.4%; Pred. No. 4.3e-95;
 Matches 267; Conservative 79; Mismatches 194; Indels 35; Gaps 15;

OY 29 PKRTVSSGELATVVRFSQTIQDFLTLTPTGLTYGAREALFAFSEALQGA-- 86
 DB 34 PRRTIPEELSGT--RHFKGQANSTLLLEFASRLVGAAGALEFSLANDIG-DGAIK 90
 OY 87 -ISMFAVPEKTECIQKKNNOTECNFRFLQPYNASHLYVGYAFQPKCTYVNMILTF 145
 DB 91 ELHWEASPEMOSKCKQKKNNOTECFNHFRFLQRLNSTLTYACGHAFOPLCAIDAIEAF 150
 OY 146 TLEHGEFEDGKGCYDPAPKAGHLLVDGELYSATLNNLGEPIILRMGPHHSKTEY 205
 DB 151 TLP-TSEFGKCKCPYDPAKGTGLTIDGLTNRFRST-PDIRSRHP-HSLRTE 207
 OY 206 LAF-WLNEPFGVSAVYVESVSGFTGDDKYVFFERAVE-----SDCYAEQVAVARA 258
 DB 208 TFMHLNDAEFYFVLYRESKASAVGDDKXYFFTERTEGSGSFTGSRSHVAVARA 267
 OY 259 RCKCDGKAGKTLQKKNFTFLKARLACSAFKNQLYFNQIOMHTLQDTSWNTTFGVF- 317
 DB 268 RCKCDGLCKKTLQKKNFTFLKARLACSAFKNQLYFNQIOMHTLQDTSWNTTFGVF- 317
 OY 318 -QAGMDMYLAICEYQLEIQVEGPEGRKEHEEAKOKDRYTDVPSRPGSCINMHR 376
 DB 324 LSTOKKTLEBAICRYDLAEIOAVAGPRMEYODSSRRKRGEGVPERPSCITDSLR 383
 OY 377 RRGYTSLELDPNLIINFVKKRLMEEOYGRMSRBLVKKGTNTHLVADRYTGIDGAYV 436
 DB 384 SQGYNSODLPSLVLDYFKLHPLMARPVYPTGRGRLLLKRNIRYTHLGTPTTPAGPTY 443
 OY 437 TYLFITGDDMLKAVSLCPWVHLLEQLF--DQPRMSLVLSQSKLLFAGSRQIYOL 495
 DB 444 DLLEFADGWIHRKAVVLSCHMIIEFQVRESSEVELYLSLQHSLYGAPGCVIOL 503
 OY 496 PVADCIKRYSCADYLRDPCYCAVSNTSRVAVGHRG-SLLIOHMTSDSGICNLRG 554
 DB 504 PLSSSRKRSCTDCLLADPYCGMPDGRHACAAATTINATLIDIERGN-----RG 556
 OY 555 --SKIQSGXP--KNITVYAGTDLVLPCHLSNLA 585
 DB 557 CESSRDTGPPPLKTRSVLRGDDVLLPCDQPSNLA 591
 RESULT 4
 SM4G_MOUSE STANDARD; PRt: 837 AA.
 AC 09MUH7;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 4G precursor.
 GN SEM4G.
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid-10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6J;
 RX MEDLINE-99425180; PubMed-10495281;
 RA Li H., Wu D.K., Sullivan S.L.;
 RT Characterization and expression of sema4g, a novel member of the
 RL semaphorin gene family.
 Mech. Dev. 87:169-173(1999).
 CC -1- FUNCTION: MAY PLAY A ROLE IN AXON GUIDANCE.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: BRAIN, SPINAL CORD, AND SEVERAL SENSORY ORGANS
 CC AS WELL AS SPECIFIC POPULATIONS OF PROJECTION NEURONS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN DEVELOPMENT.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: Contains 1 Sema domain.
 CC -1- SIMILARITY: Contains 1 Immunoglobulin-like C2-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

CC EMBL: AF134918; AAD30541.1; -;
 DR MGD: MGI:1347047; Sema4g
 DR InterPro: IPR007110; Ig_1like.
 DR InterPro: IPR003359; Ig.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003659; Plexin-like.
 DR InterPro: IPR002165; Plexin_repeat.
 DR InterPro: IPR001627; Sema.
 DR Pfam: PF01437; PSI; 1.
 DR Pfam: PF01403; Sema; 1.
 DR SMART: SM00409; IG; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00630; Sema; 1.
 DR PROSITE: PS50835; IG_Like; 1.
 KW Signal; Transmembrane; Immunoglobulin domain; Multigene family;
 KW Neurogenesis; Developmental protein; Glycoprotein.
 FT SIGNAL 1 17
 FT CHAIN 18 837
 FT DOMAIN 18 673
 FT TRANSMEM 674 694
 FT DOMAIN 695 837
 FT DOMAIN 56 487
 FT DOMAIN 505 558
 FT DOMAIN 565 649
 FT DOMAIN 563 568
 FT DISULFID 562 773
 FT CARBOHYD 55 55
 FT CARBOHYD 111 111
 FT CARBOHYD 126 126
 FT CARBOHYD 388 388
 FT CARBOHYD 540 540
 FT CARBOHYD 598 598
 SQ SEQUENCE 837 AA; 92378 MW; 5C6E9C98A545EB1 CRC64;

Query Match 38.0%; Score 1223.5; DB 1; Length 837;
 Best Local Similarity 46.9%; Pred. No. 6.3e-95;
 Matches 268; Conservative 82; Mismatches 192; Indels 29; Gaps 16;

OY 29 PKRTVSSGELATVVRFSQTIQDFLTLTPTGLTYGAREALFAFSEALQGA 86
 DB 34 PRRTIPEELSGT--RHFKGQANSTLLLEFASRLVGAAGALEFSLANDIG 91
 OY 87 -ISMFAVPEKTECIQKKNNOTECNFRFLQPYNASHLYVGYAFQPKCTYVNMILTF 146


```

Db 92 IHWEASPEMOSKCHOKGKNNOTECFNFHFLQRLNATHTFYACGTHAFOPLCAIDAETFI 151
Qy 147 LEHGEFEDGKGCPCYPDAKAGLILVDGELYSATLNFLEGTPIILRNMGPHHSKTEYL 206
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 152 LP-TSFEKGEKCEPCYPDAKAGLILVDGELYSATLNFLEGTPIILRNMGPHHSKTEYL 208
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 207 AF-WLNEPHRYGSAIYVESVSGFTGDDDKYTFEERERAVESDCAE---QVAVARAVC 261
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 209 PMHNLNAEYFVSFLVESKTSVAAGDDKLYFFEMERESESSFTQSRSHRVARAVC 268
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 262 KGDGAGATLQKRTWTFELKARLACAPNMOLYFNOLQAMPTIQ-DTSMHNTTFGVF--Q 318
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 269 KGDGAGATLQKRTWTFELKARLACAPNMOLYFNOLQAMPTIQ-DTSMHNTTFGVF--Q 323
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 319 AQMGDTLSAICEYOLEEIOYVFEKPYKEHQAOKMDRYDVPVSPRGSCINMMHRRH 378
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 324 TQMTLEASACIRDLAEIQAVFTGPMPEYDQGARCMRGYRGVPEPAPGSCITDSLSR 383
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 379 GYTSLSLPLDNLNLFVKKHPLMEQVGRPMRSPILLYKGTFTLHVDRTGLDQATTV 438
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 384 GYNSQDLPLSLVDFVKKLPMARVPYTRGRPLILKRNVTYHGTHTVSTPAGPTVDL 443
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 439 LFTGDCMLTKAVSLGPMVHLIEQLF-DQEPHRSVLTSQSKLIFAGRSQVLQPLV 497
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 444 LFTGDCMLTKAVSLGPMVHLIEQLF-DQEPHRSVLTSQSKLIFAGRSQVLQPLV 503
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 498 ADCIKYRSCADCVLARDPYCAMSNTSRCAVAGHGFSLLIQWMTSDTSGICNLRSK 556
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 504 SSCRYGSCYCIILARDPYCGMSSIHACVATVYANTELIDOI---ERNGNCEGR 559
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 557 IQSGPXP-KNITVAGTDLVLPCHLSSNLA 585
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 560 -DAGPPPLKTRSVLRGDDVLLPCDQPSNLA 589
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
SM4D_MOUSE STANDARD: PRT: 861 AA.
ID SM4D_MOUSE STANDARD: PRT: 861 AA.
AC 009126;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 4D precursor (Semaphorin J) (Sema J) (Semaphorin C-like 2)
DE (M-Sema G).
GN SEMA4D OR SEMA4 OR SEMA12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=97125976; PubMed=8969198;
RA Furuyama T., Inagaki S., Kosugi A., Noda S., Saitoh S.-I., Ogata M.,
RA Iwahashi T., Miyazaki N., Hamada T., Tohyama M.;
RA Identification of a novel transmembrane semaphorin expressed on
RA lymphocytes.;
RL J. Biol. Chem. 271:33376-33381(1996).
CC -1- FUNCTION: MAY PLAY A FUNCTIONAL ROLE IN THE IMMUNE SYSTEM, AS WELL
CC AS IN THE NERVOUS SYSTEM.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN LYMPHOID TISSUES,
CC ESPECIALLY IN THE THYMUS, AS WELL AS IN THE NERVOUS TISSUES.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: Contains 1 Sema domain.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U69535; AAC52964.1; -.
DR MGD; MGI:109244; Sema4d.
DR InterPro: IPR002447; Sema4d.
DR InterPro: IPR003598; Ig-like.
DR InterPro: IPR003006; Ig-C2.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR002165; Plexin-repeat.
DR InterPro: IPR001627; Sema.
DR Pfam: PF00047; Ig_1.
DR Pfam: PF01437; PSI_1.
DR Pfam: PF01403; Sema_1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS00835; IGC_LIKE; 1.
KV Signal: Transmembrane, Immunoglobulin domain, Multigene family;
KV Neurogenesis; Developmental protein; glycoprotein.
KV CHAIN 1
KV SIGNAL 23
KV DOMAIN 24 861
KV TRANSMEM 24 733
KV DOMAIN 734 754
KV DOMAIN 755 861
KV DOMAIN 502 551
KV DOMAIN 552 555
KV DISULF 576 624
KV CARBOHYD 49 49
KV CARBOHYD 77 77
KV CARBOHYD 139 139
KV CARBOHYD 191 191
KV CARBOHYD 379 379
KV CARBOHYD 419 419
KV CARBOHYD 613 613
KV CARBOHYD 632 632
KV SEQUENCE 861 AA; 95714 MW; 533CD6D271AD679B CRC64;

Query Match 38.0%; Score 1222.5; DB 1; Length 861;
Best Local Similarity 47.2%; Pred. No. 8e-95;
Matches 268; Conservative 82; Mismatches 197; Indels 21; Gaps 12;

Qy 28 VPKRTVSSGELATVYRFSQGTIDFLTLTLEPTGLLYGAREALEAF-SMALELOGA 86
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 27 VPKRTVSSGELATVYRFSQGTIDFLTLTLEPTGLLYGAREALEAF-SMALELOGA 84
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 87 ISWEAPVAKTEKCEOKGKNNOTECFNFHFLQRLNATHTFYACGTHAFOPLCAIDAETFI 146
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 85 ISWEAPVAKTEKCEOKGKNNOTECFNFHFLQRLNATHTFYACGTHAFOPLCAIDAETFI 144
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 147 LEHGEFEDGKGCPCYPDAKAGLILVDGELYSATLNFLEGTPIILRNMGPHHSKTEYL 206
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 145 LEHGEFEDGKGCPCYPDAKAGLILVDGELYSATLNFLEGTPIILRNMGPHHSKTEYL 202
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 207 AF-WLNEPHRYGSAIYVESVSGFTGDDDKYTFEERERAVESDCAE---QVAVARAVC 266
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 203 AF-WLNEPHRYGSAIYVESVSGFTGDDDKYTFEERERAVESDCAE---QVAVARAVC 262
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 267 GARTLQKRTWTFELKARLACAPNMOLYFNOLQAMPTIQ-DTSMHNTTFGVF--Q 326
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 263 GARTLQKRTWTFELKARLACAPNMOLYFNOLQAMPTIQ-DTSMHNTTFGVF--Q 322
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 327 SAICYOLEEIOYVFEKPYKEHQAOKMDRYDVPVSPRGSCINMMHRRHGYTS 382
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 323 SAICYOLEEIOYVFEKPYKEHQAOKMDRYDVPVSPRGSCINMMHRRHGYTS 382
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 383 SLNLPDNLNLFVKKHPLMEQVGRPMRSPILLYKGTFTLHVDRTGLDQATTVLEIG 442
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 383 SLNLPDNLNLFVKKHPLMEQVGRPMRSPILLYKGTFTLHVDRTGLDQATTVLEIG 442
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 443 TQMTLEASACIRDLAEIQAVFTGPMPEYDQGARCMRGYRGVPEPAPGSCITDSLSR 499
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 443 TQMTLEASACIRDLAEIQAVFTGPMPEYDQGARCMRGYRGVPEPAPGSCITDSLSR 502
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

QY 500 CIKYNSCADCVLARDPYCAMSVNTSRCAVVGHFGL--LIGHVMTSDTSGICNLGSKRI 557
DB 503 CEKHSCECDVCLARDPYCAMSPAIRKACVYLHQBESSRGMID-MSGDTSS-CLDKSKRS 560
QY 558 QSGPKPKNTTVVAGTDLVLPCHLSSNLA 585
DB 561 FNOHFFKHKH-----GGTAELCKFCOKSNLA 583

RESULT 6
SMAD_HUMAN STANDARD: PRT; 862 AA.
ID SMAD_HUMAN STANDARD: PRT; 862 AA.
AC 092834;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Sema4D precursor (Leukocyte activation antigen CD100) (BB18)
DE (A8) (GR3).
GN SEMA4D OR CD100.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell;
RA MEDLINE=97030273; PubMed=8876214;
RA Hall K.T., Boumsell L., Schultze V.A., Dorfman D.M.,
RA Cardoso A.A., Benussan A., Nadler L.M., Freeman G.J.;
RT "Human CD100, a novel leukocyte hemaphorin that promotes B-cell
RT aggregation and differentiation." J. Biol. Chem. 268:11780-11785 (1993).
RT Proc. Natl. Acad. Sci. U.S.A. 93:11780-11785 (1996).
CC -1- FUNCTION: MAY PLAY A FUNCTIONAL ROLE IN THE IMMUNE SYSTEM, AS WELL
CC AS IN THE NERVOUS SYSTEM. INDUCES B CELLS TO AGGREGATE AND
CC IMPROVES THEIR VIABILITY IN VITRO.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN SKELETAL MUSCLE,
CC PERIPHERAL BLOOD LYMPHOCYTES, SPLEEN, AND THYMOS AND ALSO
CC EXPRESSED AT LOWER LEVELS IN TESTES, BRAIN, KIDNEY, SMALL
CC INTESTINE, PROSTATE, HEART, PLACENTA, LUNG, AND PANCREAS BUT NOT
CC IN COLON OR LIVER.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: Contains 1 Sema domain.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD100 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd100.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U60800; AAC50810.1; -
CC Genew: HGNC:10732; SEMAD.
DR MIM: 601866;
DR GO: GO:0006916; P:anti-apoptosis; TAS.
DR GO: GO:0007155; P:cell adhesion; TAS.
DR GO: GO:0006955; P:immune response; TAS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig-MHC.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR002165; Plexin-repeat.
DR InterPro: IPR001627; Sema.
DR Pfam: PF00047; Ig; 1.
DR Pfam: PF01437; PSI; 1.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00409; IG; 1.

DR SMART: SM00423; PSI; 1.
DR SMART: SM00630; Sema; 1.
DR PROSITE: PS50835; IG-Like; 1.
KW Signal: Transmembrane; Immunoglobulin domain; Multigene family;
KW Neurogenesis; Developmental protein; glycoprotein.
FT SIGNAL 1 21
FT CHAIN 22 862
FT DOMAIN 22 734
FT TRANSMEM 735 755
FT DOMAIN 756 862
FT DOMAIN 50 482
FT DOMAIN 502 551
FT DOMAIN 554 636
FT DISULFID 576 624
FT CARBOHYD 49 49
FT CARBOHYD 77 77
FT CARBOHYD 139 139
FT CARBOHYD 191 191
FT CARBOHYD 329 329
FT CARBOHYD 379 379
FT CARBOHYD 419 419
FT CARBOHYD 613 613
FT CARBOHYD 632 632
SO SEQUENCE 862 AA; 96149 MW; 7B18FBA98789371 CRC64;

Query Match 37.9%; Score 1219.5; DB 1; Length 862;
Best Local Similarity 46.0%; Pred. No. 1.4e-94;
Matches 264; Conservative 88; Mismatches 189; Indels 33; Gaps 14;

QY 28 VPKRYTSSELTAVYRFSOTGODPLTLTPTFLVYGAREALFAR-SKALALQGA 86
DB 27 IPIITWEHREVLV--QFHEPDLYNSALLSKDCTLYIGAEVAVNALISKQHE 84
QY 87 ISWEAFVEKTEKTECKGNKQTECFNFIPLQPNASHLYVCGTVAFOPKCTVNNLTFT 146
DB 85 YVWKVSEDKKAKCAEKGKQKQECNLYINVLQSLTSLVCGTNAFOPACHLNLTSRK 144
QY 147 LEHGEFEQKCKCPDPAKAGLVDGLYSATLNNLTGTEPIILRNKPHHSKTEYL 206
DB 145 F-LGKNEKDKGRCPDPASYSYVAVDGLYSTSNFSGSEPIISRN--SSHSPLTEYA 202
QY 207 APFLNPRHFGVSAAYVESGFTGDDDKYFFEPREAVSDCAEYAVNARVAVCGDKG 266
DB 203 IPIPLNPSFVEADVIRKSPDSPDGEDRVYFEFEVSEYFVRLIPRIAVNCKGDG 262
QY 267 GARTLQKRTTFLKARLACAPNMQLYFNQLQAMHTLQDSMHTTFEFGVFOAQMGDMYL 326
DB 263 GARTLQKRTTFLKARLACAPNMQLYFNQLQAMHTLQDSMHTTFEFGVFOAQMGDMYL 322
QY 327 SAICEYQLEIQRVF-EGPYKE---YHEAQKWDRTYDVPSPRSCINNMHRRHGYTS 382
DB 323 SAVCAVNLSTAEVPSFHGKYMOSTVTEQSHTKVVRNGVPPKRPACIDSEARANYTS 382
QY 383 SLELPNLIINFVKIHPLEMEQYGRMSRPLVAKGNFTHLVADRYTGADATYTLVFLTG 442
DB 383 SLELPNLIINFVKIHPLEMEQYGRMSRPLVAKGNFTHLVADRYTGADATYTLVFLTG 442
QY 443 TGDGWLKAVSLGPAWHLTLEQLF--DQEPMSLYLSQSK--KLTFAGRSQSLVQLPVD 499
DB 443 TDRGALHKAISLEHNVHIIIEFTQQLFQDFEPVQTLILSSKKGRNFVAGSNGVQAFLAF 502
QY 500 CIKYNSCADCVLARDPYCAMSVNTSRCAVVGHFGL--LIGHVMTSDTSGICNLGSKRI 557
DB 503 CEKHSCECDVCLARDPYCAMSPAIRKACVYLHQBESSRGMID-MSGDTSS-CLDKSKRS 560
QY 552 LRGSKIOSGPKPKNTTVVAGTDLVLPCHLSSNLA 585
DB 555 DK-----SKGSYRQHFHKGGT-AELKCSQKSNLA 583

RESULT 7
SMAB_MOUSE
ID SMAB_MOUSE STANDARD: PRT; 782 AA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: AF073289; AAC7345.1; -
 DR ZFIN: ZDB-GENE-990715-7; sema7.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003659; Plexin-like.
 DR InterPro: IPR002165; Plexin-repeat.
 DR InterPro: IPR001627; Sema.
 DR Pfam: PF00047; Ig 1.
 DR Pfam: PF01437; PSI 1.
 DR Pfam: PF01403; Sema; 1.
 DR SMART: SM00409; Ig 1.
 DR SMART: SM00423; PSI 1.
 DR SMART: SM00630; Sema; 1.
 DR PROSITE: PS50835; IG_LIKE; 1.
 DR Signal: Transmembrane; Immunoglobulin domain; Multigene family;
 KM Glycoprotein.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 766 SEMAPHORIN 27.
 FT DOMAIN 25 664 EXTRACELLULAR (POTENTIAL).
 FT TRAMEM 665 685 POTENTIAL.
 FT DOMAIN 686 786 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 33 483 SEMA.
 FT DOMAIN 501 552 PSI.
 FT DOMAIN 555 640 IG-LIKE C2-TYPE.
 FT DISULFID 577 623 BY SIMILARITY.
 FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 766 AA; 85617 MW; 0ACBC693FE7D830C CRC64;

Query Match 35.6%; Score 1146.5; DB 1; Length 766;
 Query Local Similarity 42.0%; Pred. No. 17e-88;
 Matches 248; Conservative 101; Mismatches 212; Indels 29; Gaps 14;

7 WLLAARLMLGIGAEVWNNLVPRKTVSSGELATVVRFSOTGIODFTLTTEPTGLLY 66
 13 WSPAMPLGGLGSLTD---SLPKRTVPIG--SNGGRLEFRGEGINVTMLRLDNLILI 66
 67 VGARALALRFSEKALELQGA-ISMENAPVEKTECIQKGNKNOTCEPFIIRLYQYNASHL 125
 67 LGARALFALDLDITIKKAMLMKMEVTRDQNDCSNKGADATNDCKNYIRIHKKNDGRM 126
 126 YVCGATVAPKPTCYVNNML--FTFLHGEFEDEGKGCPCYDPAAGHAGLAVDELGSLATLNN 183
 127 YVCGKARNPCTGCTSYADGKLTLEILQ-EDTKGACPPDPFRRTYSAMYGAVSATSKN 185
 184 FLGTEPILRNKGRPHSMKTEFLAFLNEPHEFVGSAYVESGFTGDDKDYFFERERA 243
 186 FRGSGPVMWRS--TESIRIETFTSWLSPEFNTHAHIPEGSDNPDGDDKDYLFESERA 243
 244 VESDCYAAQVAVARAVCKDGMGARTLQKRTTTLAKARLACSDAPNMOLYENQLOAMHTL 303
 244 VEYESYTYADVSRVAVRCKGDLGSGORTLQKKTSEFLKRLDQVNTNPLLYQDVFNILC 303
 304 ODTSHNNTFPGVFEAOMGDMYLSAICYELEIORV-EGRYK-EYHREAO--FWDRYT 359
 304 PD-DFTTCVFAVFPQSDSSQYSAVCSKIEDITIVSKGKAPFVNESEFVWWMIS 362
 360 DPVPSRPGSCINNNHRRHGYSSLELFDNLINFKKPLMEEOVGPRMSRPLVKKGTN 419
 363 GELPDRPGACIDNARREKGITKLELFDKTLQPKDKPLMDQAVTA--EDELVAKRGA 420

QY 420 FTHLVADRYTGDGATYVYVLTIGTGDGWLKAVSLGPMWVHLEIQLFD-QEPMSSLYLS 478
 DB 421 FRIIVYATATATLNSHQVWFIKTSKSVLAKAVNNGEVMIEIQLDFPSPIKILRLS 480
 QY 479 OSKTLFPGSRGQVLPVADICIXRSCADYADPCYAMSVNTRSCYAVGCH---CS 535
 DB 481 SKRKDLVYGEVGVVQLSTSECGRYOTCLDYADLPCHGMDLDFEHCCTINSIRTRSS 540
 QY 536 LLIQHWMTSDTSGICNLNRSKIQSGPKNTIVVATDVLPCRHSSNTA 585
 DB 541 TVIQSLNGBDASQCPAIGVSK-----PVNISFYHGNTRYKLCGQCYRSNLA 584

RESULT 9
 SM4A_HUMAN
 ID SM4A_HUMAN STANDARD: PRT; 761 AA.
 AC 09H3SL; OSMDA9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 4A precursor (Semaphorin B) (Sema B).
 GN SEMA4A OR SEMB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RA Seki N., Hattori A., Hayashi A., Kozuma S., Muramatsu M., Miyajima N.,
 RA Saito T.;
 RT "Human semaphorin B.";
 RN Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
 RC [2]
 RC SEQUENCE FROM N.A.
 RC TISSUE=colon;
 RC MEDLINE=22386257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Scheffen C.M., Schuler G.D.,
 RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Usdin T.B., Yoshlyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
 RA Blackesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grilwood J., Schmutz J., Myers R.M.,
 RA Butlerfield V.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
 CC SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

EMBL: AB029394; BAB20087.1; -


```

OY 87 ISHEAFVEKTEKTEIOGKANNOTECENFIRFLOPYNASHLYVGTAFORPKCTVYNNLTFP 146
DB 105 IIMMPPEIRHONCRKRGK-KEDCHNFIOILAVNASHLTCTFADPCGVIDVSPQ 163
OY 147 LEHGEFEDGKGCOPYPAKGHAGLVDGELYSATLNFLETETIIRNNPHHS-KTKE 205
DB 164 -OVERLESERGRGKCFEPAPORSAVAMAGSVLTVATVNFLETETISRAGRAEDWRTET 222
OY 206 LAFWLNPEHFGVSAVY-PESVGSFTGDDVYFEFFERBAVESPCYAEQVAVARVYKGD 264
DB 223 LSWMLNAPAFVAMVLSPEMGEDEGDE-IEFFETSTRVLDSTYRINVRARACAD 281
OY 265 MCGARLQKRTTTLKARLACAPNMQLYFNOLAMHTLQDTSWNT-TFFGVFOAKGD 323
DB 282 LGGRTKLOKRTTTLKADLCCPGPEHGRASGVQANMAELRPOGACTPFFYGIFSSOWMG 341
OY 324 MYLSACEYQLEIEIOVFEGPYKEYHEBAOKMDRYD-PVSPRPGSCINNNRBRGYS 382
DB 342 MAISAVCAFRODIDRVLNGLFRELKHDCNRCLPVNDNEVPOPRPCECTANNKLOQFGS 401
OY 383 SLELPDNLINFYKHPHMEQVSPMSRDLVYKGTFTHLVADRYTGLDGYATVYLTIG 442
DB 402 SLELPDNLINFYKHPHMEQVSPMSRDLVYKGTFTHLVADRYTGLDGYATVYLTIG 461
OY 443 TGDCHILKAVSLGPRVHLIEQLF-DOEPMSLVLSOSKILLFASRSQVLQVLPVADCI 501
DB 462 TEGHILHRAVRIGAOVSLVEDLALFEPPQVSKMLYD-WLVGSHTEVQVNTSGS 519
OY 502 KYRSCADVLADPYCAMSNTSVCYA-VGSHFGSLIOHVMSTDSGICNLGSRKISG 560
DB 520 RIQSGECILQDDPYCANSFRIDACVAHNGEHRG-MVQDIESADVSLC-----PKRG 572
OY 561 PAPKNTTVAGT-DIVLPCHLSS 582
DB 573 EHPVVEFVPAVATGVHVLPCSPSS 596

```

RESULT 11

SMAF_HUMAN STANDARD; PRT; 770 AA.

AC 095754; Q9NSJ5;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Semaphorin 4F precursor (Semaphorin W) (Sema W) (Semaphorin M) (Sema M).

GN SEMA4F OR SEMA W OR SEMA W.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_TaxID=9606;

OK

RN

RP

RP SEQUENCE FROM N.A. (ISOFORM SMALL).

RC TISSUE=Brain;

RX MEDLINE=9916263; PubMed=10051670;

RA Encinas J.A., Kikuchi K., Chedotal A., de Castro F., Goodman C.S., Kimura T., Cloning, expression, and genetic mapping of Sema W, a member of the semaphorin family.

RT Proc. Natl. Acad. Sci. U.S.A. 96:2491-2496(1999).

RT [2]

RN

RP

RP SEQUENCE FROM N.A. (ISOFORM LONG).

RA Chen F., Do A., Do T., Weisler M., Roe B.A.; Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP

RP TISSUE=amygdala;

RC

RC SEQUENCE FROM N.A. (ISOFORM LONG).

RX MEDLINE=2154917; PubMed=11230166;

RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S., Ansgore W., Boecker M., Bloecher H., Bauersachs S., Blum H., Lauber J., Duestenhoef A., Beyer A., Koehler K., Strack N., Mewes H.-W., Othenwelder B., Obermaier B., Tampe J., Heubner D., Wandt R., Korn B., Klein M., Poustka A.;

```

RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs."
RT Genome Res. 11:422-435(2001).
RN [4]
RA Jang W., Spillson S.V., Hua A., Roe B., Weisler M.H.;
RT Large-scale comparative sequence analysis of human and mouse genomic
RT DNA in the mnd2 region of mouse chromosome 6 reveals coding regions of
RT three new genes."
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: HAS GROWTH CONE COLLAPSE ACTIVITY AGAINST RETINAL
CC GANGLION-CELL AXONS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-alternative splicing; Named isoforms=2;
CC Name=long;
CC IsoId=Q95754-1; Sequence=Displayed;
CC Name=short;
CC IsoId=Q95754-2; Sequence=VSP_006043;
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: Contains 1 Sema domain.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB022317; BAA75631.1; -
CC DR EMBL: AC006544; -; NOT_ANNOTATED_CDS.
CC DR EMBL: AL136552; CAB66487.1; -
CC DR EMBL: AF053369; AAF80660.1; -
CC DR GeneW: HGNC:10734; SEMA4F.
CC DR MIM: 603706; -
CC DR GO: GO:0005887; C: integral to plasma membrane; TAS.
CC DR GO: GO:0005624; C: membrane fraction; TAS.
CC DR GO: GO:0007411; P: axon guidance; TAS.
CC DR GO: GO:0007267; P: cell-cell signalling; TAS.
CC DR GO: GO:0007399; P: neurogenesis; TAS.
CC DR InterPro: IPR003659; Plexin-like.
CC DR InterPro: IPR002165; Plexin_repeat.
CC DR Pfam: PF01437; PSI; 1.
CC DR Pfam: PF01403; Sema; 1.
CC DR SMART: SM00423; PSI; 1.
CC DR SMART: SM00630; Sema; 1.
CC Signal: Transmembrane; Immunoglobulin domain; Multigene family;
CC Neurogenesis; Developmental protein; Glycoprotein;
CC KW Alternative splicing.
CC FT SIGNAL 1 34
CC FT CHAIN 35 770
CC FT DOMAIN 35 659
CC FT TRANSMEM 660 680
CC FT DOMAIN 681 770
CC FT DOMAIN 65 496
CC FT DOMAIN 512 563
CC FT DOMAIN 580 635
CC FT DISULFID 587 628
CC FT CARBOHYD 64 64
CC FT CARBOHYD 133 133
CC FT CARBOHYD 509 509
CC FT VARSPLIC 120 274
CC FT CONFLICT 533 533
CC FT SEQUENCE 770 AA; 83511 MW; CFB874B410F0E9C8 CRC64;

```

Query Match 31.2%; Score 1005.5; DB 1; Length 770;

Best local similarity 39.4%; Pred. No. 1.3e-76;

Matches 222; Conservative 100; Mismatches 222; Indels 19; Gaps 15;


```

OY 28 VPRVTSSGELATVVRPSQTGIDFLTLTEPTGLLYVGAREALFAFSME-ALELOGA 86
DB 40 VPRVSLPSEADSLTREFAVPHTYNSLYLVDPAHSHTLYVGARPTIFALSLPFSEGERPR 99
OY 87 ISWAPVEKTECJOKGNKNOTECGFNFREFLOPYNASHLYVCGTYAPRCKTYVMMLFT 146
DB 100 IDMWVPEAHNRCKRKGK-KEDECHNFVOLIILANASHLLTCGTFAPFCGVIDVSRFO 158
OY 147 LEHGEFEDGKGCPCYDPKAGHAGLYVDELVSATLNNFLGTEPTILRNNGPHNS-MKTEY 205
DB 159 -OVERLESGRGKCFEPERORSAAMAGVLAATYKNTLGPETITRAVGAEDMIRDTI 217
OY 206 LAFMLNEPHTVGS-AVPESVSGTGDGDKYFFFRERAVESDCYAEQVAVARAVCKGD 264
DB 218 LPSMLNAPAFVAAVALSPAEWGDEGODE-ITFEFTETSRADFSTERYIKVPRVAVACAD 276
OY 265 MCGARTLOKRTTFLKARLACASAPNMQLYFNOLQAMHTLQDTSKINT-TFEGVQAOQGD 323
DB 277 LGGKRTLOORWTLTKADLLCPGPHGRASSVLQDVAALREBLGAGTPIFYGIFSSOWEG 336
OY 324 MYLSAICEYOLEEIORVEGPGYKEYHEBAOKMDRTD-PVPSRPGSCIT-NMHRHRYT 381
DB 337 ATISAVCAFRPDIDRTYVNGFRELKHCNGLPVVDVQPRGECITNMKLIRH-FG 395
OY 382 SLELPMNLINFKKHPMEBOVGRMSRPLVKKGTNFTLVADRTVGLDGAITYVLEFI 441
DB 396 SLSLDPDVLTFIRHIDRPLMDRPFADGRLPLVTTDTAYLRVAVHRTVSLSGKEYDVLYL 455
OY 442 GTGGWMLKAVSLGFWHLIEQLF-DQEPKRSVLVSOSKLEFAGRSQVLQVLPVADC 500
DB 456 GTEGHILRAVRIGAOQLSVLEDLALFPPRPVENNKLKHS--WLVGSRTEVTOVNTNGC 513
OY 501 IKYRSCADCVLADBPYCAVSNTSKVCA-VGHEFSGSLTIQHWISDTSGICNLKSGKIOS 559
DB 514 GRGSCSCLIAQDPYCAMSFRLDECVAHAGEHRG-LVQIDESADVSLSCKEKG--- 568
OY 560 GPXRNITVAGTDLVLPCHLSS 582
DB 569 RPVVEFVAVARAHVLPSPSS 591

```

RESULT 12

SK4F_MOUSE STANDARD: PRF: 777 AA.

ID SK4F_MOUSE 092123; 09R1Y1; 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Semaphorin 4F precursor (Semaphorin W) (Sema W).

GN SEMA4F OR SEMA4.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

NCBI_Taxid=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALE/c; and 129/Svt; TISSUE=Brain;

RX MEDLINE=99162633; PubMed=10051670;

RA Enclina J.A., Kituchi K., Chedotal A., de Castro F., Goodman C.S., Kimura T.

RT "Cloning, expression, and genetic mapping of Sema W, a member of the semaphorin family."

RL Proc. Natl. Acad. Sci. U.S.A. 96:2491-2496(1999).

CC -1- FUNCTION: HAS GROWTH CONE COLLAPSE ACTIVITY AGAINST RETINAL GANGLION-CELL AXONS (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.

CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib.ch).

```

CC EMBL; AB021291; BAA75630.1; -.
CC EMBL; AB022316; BAA75634.1; -.
CC EMBL; AB022312; BAA75634.1; JOINED.
CC EMBL; AB022312; BAA75634.1; JOINED.
CC EMBL; AB022312; BAA75634.1; JOINED.
CC EMBL; AB022314; BAA75634.1; JOINED.
CC EMBL; AB022315; BAA75634.1; JOINED.
CC MGI; MGI:1340055; Sema4f.
CC InterPro: IPR003659; Plexin-like.
CC InterPro: IPR002165; Plexin_repeat.
CC InterPro: IPR001627; Sema.
CC Pfam; PF01437; PSI; 1.
CC Pfam; PF01403; Sema; 1.
CC SMART; SM00423; PSI; 1.
CC SMART; SM00630; Sema; 1.
CC Signal; Transmembrane; Immunoglobulin domain; Multigene family; Neurogenesis; Developmental protein; Glycoprotein.
CC SIGNAL 1 40 POTENTIAL.
CC CHAIN 41 777 SEMAPHORIN 4F.
CC DOMAIN 41 667 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 668 688 POTENTIAL.
CC DOMAIN 689 777 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 71 502 SEMA.
CC DOMAIN 518 569 PSI.
CC DOMAIN 586 641 IG-LIKE C2-TYPE.
CC DISULFID 593 634 BY SIMILARITY.
CC CARBOHYD 70 70 N-LINKED (GLCNAc. . .) (POTENTIAL).
CC CARBOHYD 139 139 N-LINKED (GLCNAc. . .) (POTENTIAL).
CC CARBOHYD 515 515 N-LINKED (GLCNAc. . .) (POTENTIAL).
CC VARIANT 490 490 P -> T (IN STRAIN 129/Svt).
CC VARIANT 659 659 S -> A (IN STRAIN 129/Svt).
CC SEQUENCE 777 AA; 84501 MW; B53F85358585924 CRC64;

```

Query Match 31.1%; Score 1000.5; DB 1; Length 777;

Best local similarity 39.2%; Pred. No. 3.4e-76;

Matches 221; Conservative 102; Mismatches 220; Indels 21; Gaps 14;

```

OY 28 VPRVTSSGELATVVRPSQTGIDFLTLTEPTGLLYVGAREALFAFSME-ALELOGA 86
DB 40 VPRVSLPSEADSLTREFAVPHTYNSLYLVDPAHSHTLYVGARPTIFALSLPFSEGERPR 105
OY 87 ISWAPVEKTECJOKGNKNOTECGFNFREFLOPYNASHLYVCGTYAPRCKTYVMMLFT 146
DB 100 IDMWVPEAHNRCKRKGK-KEDECHNFVOLIILANASHLLTCGTFAPFCGVIDVSRFO 164
OY 147 LEHGEFEDGKGCPCYDPKAGHAGLYVDELVSATLNNFLGTEPTILRNNGPHNS-MKTEY 205
DB 159 -OVERLESGRGKCFEPERORSAAMAGVLAATYKNTLGPETITRAVGAEDMIRDTI 223
OY 206 LAFMLNEPHTVGS-AVPESVSGTGDGDKYFFFRERAVESDCYAEQVAVARAVCKGD 264
DB 218 LPSMLNAPAFVAAVALSPAEWGDEGODE-ITFEFTETSRADFSTERYIKVPRVAVACAD 282
OY 265 MCGARTLOKRTTFLKARLACASAPNMQLYFNOLQAMHTLQDTSKINT-TFEGVQAOQGD 323
DB 277 LGGKRTLOORWTLTKADLLCPGPHGRASSVLQDVAALREBLGAGTPIFYGIFSSOWEG 342
OY 324 MYLSAICEYOLEEIORVEGPGYKEYHEBAOKMDRTD-PVPSRPGSCIT-NMHRHRYT 382
DB 337 ATISAVCAFRPDIDRTYVNGFRELKHCNGLPVVDVQPRGECITNMKLIRH-FG 402
OY 382 SLELPMNLINFKKHPMEBOVGRMSRPLVKKGTNFTLVADRTVGLDGAITYVLEFI 442
DB 396 SLSLDPDVLTFIRHIDRPLMDRPFADGRLPLVTTDTAYLRVAVHRTVSLSGKEYDVLYG 462
OY 442 GTGGWMLKAVSLGFWHLIEQLF-DQEPKRSVLVSOSKLEFAGRSQVLQVLPVADC 501
DB 456 GTEGHILRAVRIGAOQLSVLEDLALFPPRPVENNKLKHS--WLVGSRTEVTOVNTNGC 520

```


CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- DEVELOPMENTAL STAGE: At E-11, expression was restricted to the
 CC OLFACATORY PIT, THE BASAL AND NOSTRAL SURFACE OF THE TELENEPHALIC
 CC VESICLE, THE EYE ANLAGE, THE EPITHELIUM OF RATHKE'S POUCH, AND
 CC SOMITES. AT LATER DEVELOPMENTAL STAGES, IT WAS WIDELY DISTRIBUTED
 CC IN NEURONAL AS WELL AS IN MESENCHYMAL AND EPITHELIAL STRUCTURES
 CC OUTSIDE THE NERVOUS SYSTEM. AFTER BIRTH, MESENCHYMAL LEVELS
 CC DECREASED RAPIDLY AND EXPRESSION BECAME RESTRICTED TO SPECIFIC
 CC SETS OF NEURONS IN THE CNS. IN THE MATURE CNS, IT IS DETECTABLE IN
 CC MITRAL CELLS, NEURONS OF THE ACCESSORY BULB AND CEREBRAL CORTEX,
 CC CEREBELLAR PURKINJE CELLS, AS WELL AS A SUBSET OF CRANIAL AND
 CC SPINAL MOTONEURONS.
 CC -1- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
 CC THIRD OF THE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: Contains 1 semaphorin-like C2-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X55286; CAA64607.1; -
 CC InterPro: IPR007110; Ig-like.
 CC InterPro: IPR003599; Ig.
 CC InterPro: IPR003006; Ig_MHC.
 CC InterPro: IPR003659; Plexin-like.
 CC InterPro: IPR001627; Sema.
 CC Pfam: PF00047; Ig 1.
 CC Pfam: PF01403; Sema; 1.
 CC SMART: SM00409; Ig; 1.
 CC SMART: SM00423; PSI; 1.
 CC SMART: SM00630; Sema; 1.
 CC PROSITE: PS50835; IG-LIKE; 1.
 CC Signal: Immunoglobulin domain; Multigene family; Neurogenesis;
 CC developmental protein; Glycoprotein.
 CC SIGNAL 1 20 POTENTIAL.
 CC FT CHAIN 21 772 SEMAPHORIN 3A.
 CC FT DOMAIN 240 538 SEMA.
 CC FT DOMAIN 577 665 IG-LIKE C2-TYPE.
 CC FT DOMAIN 728 770 ARG/LYS-RICH (BASIC).
 CC FT DISULFID 650 723 BY SIMILARITY.
 CC FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SO SEQUENCE 772 AA; 88808 MW; 240907B12FF9F2D2 CRC64;

Query Match 29.1%; Score 937.5; DB 1; Length 772;
 Best Local Similarity 38.6%; Pred. No. 6.8e-71;
 Matches 208; Conservative 95; Mismatches 199; Indels 37; Gaps 15;

QY 10 LAARIMGLIGAEVWV-----NLVPRKTVSSGELATVVRFSOTGI---ODFLTLTLEPT 62
 DB 7 IACLEFWGILLTARAVYANGKNVPRKLSYKEMLESNNVTTFNGLANGSSYHTFLDEER 66

QY 63 GLLYVGAREALFAFMEALELOGAISWEAPVEKTECIQKKNQTECFNIRFLQPVNA 122
 DB 67 SRLYVGADHITFSLVNIKDKQKIWPVSYTRRDECKKAGKDIKECANFKVLKAYNQ 126

QY 123 SHLYVCGTYARQPKCTVYNN-----LFTLEHGEFEDGKCKPYDPAKGHAGLLVDGE 176
 DB 127 THLYVACGAGAFPICTYIEVGHNPEDNFKLQDSHFENGGRGSKPYDKILTLASLLDGE 186

QY 177 YSATLNNFLGTEPIITLRNKGPRHSMKTE-YLAFWLNEPHVGSAYVPESVGSFTGDDKV 235
 DB 187 YSGTAAEDMGDRFAIFRTLGHHPIRTQHDSCRMLNDPRFISAILIPESDNP---EDDKV 243

QY 236 YFFFERAVESDCYAEQVAVARVCKDGMGARTLQKRWTFELKARLACGAP---NMQL 292
 DB 444 YFFFERNAIDGESHGKATNARIQICNDPFGHRSLVNKNWTFELKARLICSVPNGIDT 303

DB 244 YFFFERNAIDGESHGKATNARIQICNDPFGHRSLVNKNWTFELKARLICSVPNGIDT 303
 QY 293 YFNQLOAMHTLIDTSMHNTTFEGVFOQWMDMLSAICEYOLEIQRVEGPKYKEHEA 352
 DB 304 HPELODVLNMSKDKPNPIYGVFTSSNIFGSAVCATSMADVRRVLGPRAHNDGR 363

QY 353 OKMDRTDPVPSRPGSCINNMHRNGYSSLELPNINLFWKHPIMEBOVGRNRP 412
 DB 364 YQWVPYQGRVYPYRPTCSKTP--GGPDSTKDLPDVITFAFASHPMANPVPINNRP 421

QY 413 LVKKGNT--FTHLVADRTVGLDGTATYVLEIGTGDMLKAVSL--GPNVH-----LIE 463
 DB 422 MITDVNYQFTQIVDVRVAEDG--QYDVFIFGIDVGTALKVSVPRETW--HDEEVLLE 479

QY 464 LQLEQDEP--MRSVLVSQSKLFLFAGSRQVQLPVADCIKY--RSCADCVLNDPYCAW 519
 DB 480 MIVF--REPTTISAMELSTQOOLYISTGTAGVQLPLHRCDIYKACACGLANDPYCAW 537

Search completed: August 7, 2003, 13:48:02
 Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2003, 13:45:14 ; Search time 105 Seconds
(without alignments)
1464.756 Million cell updates/sec

Title: US-10-002-050-14

Perfect score: 3219
Sequence: 1 MAFHMAVWLLAARLWLGIG.....PCHLSSNLAIPDSNPESSV 596

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea: *
2: SP bacteria: *
3: SP fungi: *
4: SP human: *
5: SP invertebrate: *
6: SP mammal: *
7: SP mhc: *
8: SP organelle: *
9: SP phage: *
10: SP plant: *
11: SP rodent: *
12: SP virus: *
13: SP vertebrate: *
14: SP unclassified: *
15: SP virus: *
16: SP bacteriophage: *
17: SP archaea: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2918	90.6	963	4 09C0C4	09C0C4 homo sapien
2	1353.5	42.0	510	4 09NYS2	09NYS2 homo sapien
3	1240	38.5	893	4 09C0B8	09C0B8 homo sapien
4	1179.5	36.6	823	11 08BIR6	08BIR6 mus musculu
5	1178.5	36.6	799	11 08BUC1	08BUC1 mus musculu
6	1169.5	36.3	550	11 08BIC3	08BIC3 mus musculu
7	1029.5	32.0	761	4 08WUA9	08WUA9 homo sapien
8	913	28.4	774	13 08JTW9	08JTW9 xenopus lae
9	884.5	27.5	748	4 08TB71	08TB71 homo sapien
10	878.5	27.3	777	11 08BH34	08BH34 mus musculu
11	876.5	27.2	754	4 08TDV7	08TDV7 homo sapien
12	872.5	27.1	777	11 08BMF6	08BMF6 mus musculu
13	852.5	26.5	775	11 09OXK3	09OXK3 mus musculu
14	851.5	26.5	782	4 09NS98	09NS98 homo sapien
15	818	25.4	756	13 08OCU9	08OCU9 gallus gall
16	779	24.2	635	4 096GX0	096GX0 homo sapien

17	730.5	22.7	1202	4 09P283	09P283 homo sapien
18	729	22.6	998	4 08NFY6	08NFY6 homo sapien
19	729	22.6	1011	4 08NFY3	08NFY3 homo sapien
20	729	22.6	1017	4 08NFY5	08NFY5 homo sapien
21	729	22.6	1022	4 09P249	09P249 homo sapien
22	729	22.6	1073	4 08NFY4	08NFY4 homo sapien
23	727	22.6	595	11 08BK68	08BK68 mus musculu
24	726.5	22.6	935	4 096JF8	096JF8 homo sapien
25	726	22.6	844	11 08BXU8	08BXU8 mus musculu
26	706.5	21.9	963	11 091X36	091X36 mus musculu
27	693	21.5	1049	4 09P2H9	09P2H9 homo sapien
28	686	21.3	1030	4 09H2E6	09H2E6 homo sapien
29	683	21.2	697	5 08MLP1	08MLP1 drosophila
30	681	21.2	724	5 09V7Q7	09V7Q7 drosophila
31	672.5	20.9	605	11 08BXZ7	08BXZ7 mus musculu
32	672	20.9	587	11 08BUT0	08BUT0 mus musculu
33	646	20.1	1005	11 09E071	09E071 mus musculu
34	645.5	20.1	476	4 08NFY7	08NFY7 homo sapien
35	645	20.0	403	11 08BK06	08BK06 mus musculu
36	641.5	19.9	687	4 09BXR8	09BXR8 homo sapien
37	637	19.8	202	13 08UYD5	08UYD5 xenopus lae
38	635	19.7	770	5 044253	044253 drosophila
39	634	19.7	770	5 09V3K4	09V3K4 drosophila
40	633.5	19.7	616	5 09V7P8	09V7P8 drosophila
41	615	19.1	457	4 09HBR1	09HBR1 homo sapien
42	614	19.1	367	4 09HAR9	09HAR9 homo sapien
43	614	19.1	418	4 096JG9	096JG9 homo sapien
44	606.5	18.8	923	11 08RAU3	08RAU3 mus musculu
45	604.5	18.8	920	11 08RAU4	08RAU4 rattus norv

ALIGNMENTS

RESULT 1
ID 09C0C4 PRELIMINARY; PRT; 963 AA.
AC 09C0C4;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical Protein KIAA1739 (Fragment).
GN KIAA1739.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RS SEQUENCE FROM N.A.
RA MEDLINE-21082932; PubMed-11214970;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:347-355(2000).
DR EMBL; AB051526; BAB21830.1; -
DR Genbank; HGNC:10731; SEMA4C.
DR InterPro; IPR003599; IG_1.
DR InterPro; IPR007110; IG_1like.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin-repeat.
DR InterPro; IPR001627; SEMA.
DR Pfam; PF01437; PSI; 1.
DR Pfam; PF01403; SEMA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; SEMA; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 963 AA: 106735 MW: 3FA37DBA3483ECE CRC64;
Query Match 90.6%; Score 2918; DB 4; Length 963;

Best Local Similarity 92.2%; Pred. No. 3,5e-253;
Matches 554; Conservative 5; Mismatches 22; Indels 20; Gaps 5;

8 W-LAALMGLOIGAEVMMNLVPRKTVSS-----GELATVVRFSQGTGIDPFLITLLE 60
DB 133 WHLVA--WGAGSRGARLRAYEPOSSCSAALMTPLAELATVRRRSQGTGIDPFLITLLE 190
61 PGLLYVAREALAFPSMEALELQCAISWEAPVEKTECIQKNNQCECFNIFLPQY 120
DB 191 PGLLYVAREALAFPSMEALELQCAISWEAPVEKTECIQKNNQCECFNIFLPQY 250
121 NASHLVGGTAAFOFKCTYVNMVLTLEHGEFEDGKGPYPAGHGGLVADGELYSAT 180
DB 251 NASHLVGGTAAFOFKCTYVNMVLTLEHGEFEDGKGPYPAGHGGLVADGELYSAT 310
181 LNPLFTEPIILRNKGPVHSMKTEYLAFMLNEPHEVGSAYVESGSGFTGDDDKYFFER 240
DB 311 LNPLFTEPIILRNKGPVHSMKTEYLAFMLNEPHEVGSAYVESGSGFTGDDDKYFFER 370
241 ERAVSDCYAEQVAVARAVKCGKMGARTLQKRTTFLKARLACSPRMQLYFNOLOM 300
DB 371 ERAVSDCYAEQVAVARAVKCGKMGARTLQKRTTFLKARLACSPRMQLYFNOLOM 430
301 HTLQDTSMHNTFFGVFOAGMDYLSAICEYQLEIQRVFEQPKYEHHEAOKMDRYTD 360
DB 431 HTLQDTSMHNTFFGVFOAGMDYLSAICEYQLEIQRVFEQPKYEHHEAOKMDRYTD 490
361 PVPSPRPGSCINNMHRHRCYTSSELDPNINLFVKHPLMEQVGRPMRSLPLVKKGNF 420
DB 491 PVPSPRPGSCINNMHRHRCYTSSELDPNINLFVKHPLMEQVGRPMRSLPLVKKGNF 550
421 THLVADRYTGDLGATYTYLFTGTGGMILKAVSLCPVWHLIEQLDFQEPKRSVLSS 480
DB 551 THLVADRYTGDLGATYTYLFTGTGGMILKAVSLCPVWHLIEQLDFQEPKRSVLSS 610
481 KLLFAGRSQSLVQVADCIKRYSCADCVLARDPYCANSVTSRCVAVAGHSGSLIOH 540
DB 611 KLLFAGRSQSLVQVADCIKRYSCADCVLARDPYCANSVTSRCVAVAGHSGSLIOH 670
541 VMTSDTSGICNLGSKIGSGPKNITVAGTDLVLPCHLSNLA-----LPDSN 590
DB 671 VMTSDTSGICNLGSKIGSGPKNITVAGTDLVLPCHLSNLAHARVTFGGRDLPAEQ 729
591 P 591
DB 730 P 730

RESULT 2
09NX92 PRELIMINARY; PRT: 510 AA.
AC 09NX92;
DB 01-OCT-2000 (TREMBLrel. 15, Created)
DB 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DB 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ20369.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Odayashi M.,
RA Nishii T., Shibahara T., Tanaka T., Nakamura Y., Isegaki T., Sugano S.,
RT NEDO human cDNA sequencing project.*;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK000376; BA01124.1; -
DR InterPro: IPR003599; Ig.-like.
DR InterPro: IPR007110; Ig.-like.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR002165; Plexin-repeat.
DR InterPro: IPR001627; Sema.

DR Pfam: PF01437; PSI: 1.
DR Pfam: PF01403; Sema: 1.
DR SMART: SM00409; IG: 1.
DR SMART: SM00423; PSI: 1.
DR PROSITE: PS00835; IG-Like: 1.
KW Hypothetical protein
SQ SEQUENCE 510 AA; 53125912E9AD0F04 CRC64;

Query Match 42.0%; Score 1353.5; DB 4; Length 510;
Best Local Similarity 92.4%; Pred. No. 6,9e-113;
Matches 257; Conservative 2; Mismatches 8; Indels 11; Gaps 2;

324 MTLAICEYQLEIQRVFEQPKYEHHEAOKMDRYDPPSPRPGSCINNMHRHRCYTS 383
DB 1 MTLAICEYQLEIQRVFEQPKYEHHEAOKMDRYDPPSPRPGSCINNMHRHRCYTS 60
384 LELPDNINLFVKHPLMEQVGRPMRSLPLVKKGNFTLHVADRYTGDLGATYTYLFT 443
DB 61 LELPDNINLFVKHPLMEQVGRPMRSLPLVKKGNFTLHVADRYTGDLGATYTYLFT 120
444 GDGWLKAVSLCPVWHLIEQLDFQEPKRSVLSSKLLFAGRSQSLVQVADCIKRY 503
DB 121 GDGWLKAVSLCPVWHLIEQLDFQEPKRSVLSSKLLFAGRSQSLVQVADCIKRY 180
504 RSCADCVLARDPYCANSVTSRCVAVAGHSGSLIOHVTSDTSGICNLGSKIGSGP 563
DB 181 RSCADCVLARDPYCANSVTSRCVAVAGHSGSLIOHVTSDTSGICNLGSKIGSGP 239
564 KNTVAGTDLVLPCHLSNLA-----LPDSN 591
DB 240 KNTVAGTDLVLPCHLSNLAHARVTFGGRDLPAEQ 277

RESULT 3
09C0B8
ID 09C0B8 PRELIMINARY; PRT: 893 AA.
AC 09C0B8;
DB 01-JUN-2001 (TREMBLrel. 17, Created)
DB 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DB 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein KIAA1745 (Fragment).
GN KIAA1745.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082932; PubMed=11214970;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.,
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.*;
RL DNA Res. 7:347-355(2000).
DR EMBL: AB051532; BAB21836.1; -
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR002165; Plexin-repeat.
DR InterPro: IPR001627; Sema.
DR Pfam: PF01437; PSI: 1.
DR Pfam: PF01403; Sema: 1.
DR SMART: SM00423; PSI: 1.
DR SMART: SM00630; Sema: 1.
KW Hypothetical protein.
FT NON-TER 1
SQ SEQUENCE 893 AA; 98361 MW; D66C48DEE524F14 CRC64;

Query Match 38.5%; Score 1240; DB 4; Length 893;
Best Local Similarity 45.4%; Pred. No. 2,5e-102;
Matches 265; Conservative 88; Mismatches 201; Indels 30; Gaps 12;

25 WNLVPRKTVSSGELATVVRFSQGTGIDPFLITLLEPGLYVAREALAFPSMEALELQ 84
DB 98 WNLVPRKTVSSGELATVVRFSQGTGIDPFLITLLEPGLYVAREALAFPSMEALELQ 157


```

0Y      85 GA-----ISEAEPKTEETECIOGKNNOCTCFNPIETLOIYVNSHSLVVCGYAPROPCITY 140
        : : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db     156 GGEYOELLMGDALEKKQOCSEFKCKDPRBOCQNTIKILPLBSGHLETFCGTAAFSMCTYT 217

0Y      141 NMLFTPLEHGE----FEDOKGCPYPAPKAHGAGLLVDGELYSATLNNFLGTEPITLRNM 195
        : : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db     218 NNENEFNLADEKGWLVLEDKGCKRCEPDPMFKSTALVDELXYGVTSFFOGNDPAISRQ 277

0Y      196 GPHHSKTEYLAFWLNEPHFGSAAYVESVSFTGDVVYFEPPREARAESDCYAQOYVA 255
        : : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db     278 SLRPT-KTSSLSLNMDPAFWASATPEBISGLSQGGDDXIYFEESTGOEFFEFPENTIVS 336

0Y      256 RVARYCGDMGCAARTLORRWTFYLKARLACSAFNMOLYENOLOAMHTLODT--SMHTTF 313
        : : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db     337 RIARIKGBEGGERVLOQRWTSFLKAOCLCSRPDDCFPNVLQDVFTLSPSPODMRDITLE 396

0Y      314 PCVFQAW--GDMYSAICEYOUELFQRYPEBGKYKHHEAQWMDTYTPVPSPRGSC 371
        : : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db     397 YCFPFSSOMRGTTGEGSAVCFMKDKDQRFVSGLYKRVNNNETQDWVTYTHVHPFRGCAC 456

0Y      372 NNWHHRNGYTSLELPDNLIINFVKKPHBLEEOVGPRMSRLYTKKGTNFTHLVADRYTGL 431
        : : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db     457 TNSARERKINSSLQLPDRVLTLEKDFLMDQGVR--SRMLTLQPQARGOVAHVHRPGL 513

0Y      432 DGAITYVLEITGODGHLKAVSLGPPVNHLEELQLDO--EPKMSVLSOSKILLFKGSR 490
        : : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db     514 H-HTYDLVFLGTGDXGLRAVSVGPVNIHIEELQITSSQOPVNLLDTHRGLLYAASH 572

0Y      491 QLVOLPVACICYYRCADCVLARDPYCANSVNTSRCVAAGHHSGL---IGWMTSDT 546
        : : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db     573 GVVOYVMANCSLYRSCGCDLLARDPYCAN--GSSCKHSLVQPOLATRPWIODIGASA 630

0Y      547 SGICN-----LRSGIOSGPXP-KNITVAGTDLVLPCHISSNIA 585
        : : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db     631 KULGSASSVYSPBFVTEKPCEQOVOPONTVWTATCPLLSNIA 674

```

```

RESULT 4
OBIR6
ID OBIR6 PRELIMINARY: PRT: 823 AA.
AC OBIR6:
DT 01-MAR-2003 (TREMBlrel, 23, Created)
DT 01-MAR-2003 (TREMBlrel, 23, last sequence update)
DT 01-MAR-2003 (TREMBlrel, 23, last annotation update)
DE Semaphorin 4B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI-TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary, and Uterus;
RX MEDLINE=22354683; PubMed=12466851;
RA THE PANTOM Consortium.
RT THE RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK030331; BAC26906.1; -.
SQ SEQUENCE 823 AA; 91485 MW; E3B8C5E5CPA1CD45 CRC64;

Query Match 36.6%; Score 1179.5; DB 11; Length 823;
Best Local Similarity 43.1%; Pred. No. 6.1e-97;
Matches 239; Conservative 92; Mismatches 215; Indels 35; Gaps 122

OY 7 VLLAARLWGLGIGAEVWNNLVPRKTVSSGELATVVRRESQTGIODFLTLTTEPTGLLY 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 16 VILLLLLLLTWTTTTRA-----LGPRIWVPLCCGGERLIRKREAMENISYTLALLSQGKILY 70

OY 67 VGAREALVAFPSMALELQQA-----ISMENPVEKTEKTEICIOGKNNOTECNEFRLOPYNA 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 71 VGARREALVAFINSLSLPGGEGYOEILMSADARKKQDCSKGDPKRCQNYIKILPLNS 130

```

```

0Y 123 SHLLVYGAJPAFQKCYVMNLFTEFLENH-----FEDGKQCPYDPAKHAAGLJNGEYL 177
Db 131 SHLLVCGTAAEPBLCAVTHIASFTLADAGAVILEDGKRCFDEFNEFSTALVYGEYL 190
0Y 178 SATLNNFLGTEDPILIRNMGRPHSHMKTEYLAFMLNEBHPFGSAVPEPVSFTGDDDDRYE 237
Db 191 TGTVSFQGNDAIPRSGSSRPLT-KVESSLNMLQDBAFVAAVPEPVSLSPTGGDDRIYE 249
0Y 238 PFREPAVEDSCAEDVYAVAVAYCKDDGAGARTLQAKWTFTEKLAKLACSPMKQLYFNOL 297
Db 250 PFSEFGQEEFEPENTIVSRVARNCKDEGGERLQOARWMSFLKAOLLCSPRDGPFFNYL 309
0Y 298 QAMHTLQD-----SMHNTTFPGVFQAOV---GDMVLSAICEQLEBTEIYAFVBEQYKEHEAO 353
Db 310 QDVFLLNPNRQDMRKLFYGVFTSQHNRGTTEBSALCEFTMNDVQAQFGLYKAVRETD 369
0Y 354 KMDRYTDPVPSRPGSCINNMRHRCYSTLSLELPDNLINFKVKKPLMEBOYGRMSRPLL 413
Db 370 QYTTEFHQVTPRPGACINRSAREKINSILOPDEVLFNLKHPFLMDGQVVR---SRLL 426
0Y 414 VKGTNFTLVADRYTGLDGAATYVLEFCTGDDPMLKAVSLGWMHLELEOLFEDP-EPH 472
Db 427 LQPRARYQVAVHRYVGLH-STYDVLFCTGGRKHLKATYLSRVIHILELOLFPBGQPV 485
0Y 473 RSLVYSQSKLLPAGRSQLOVOLPVADCTIKYNSCADCYLARDPYCAWSVNTSHCAVVGH 532
Db 486 QMILLDSHGGLLYASSHSGVQVVPVANCISLYPCGCCLLAPDYCAWTSAGACIASY--- 542
0Y 533 FCSLLIQHWYSDTSG-----ICNLNGSK-----IQSGPXRNILVVGATDLYLCHLSSN 584
Db 543 YRPDLASRWTDIDEGASVKELCNKSSTYARFLVPBCKQVOIOQPTVYVTLICPLISNL 602
0Y 585 A 585
Db 603 A 603

```

```

RESULT 5
Q8BJCI
ID Q8BJCI PRELIMINARY; PRT; 799 AA.
AC Q8BJCI;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, last annotation update)
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOJ; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA THE FANTOM Consortium,
RA "The Riken Genome Exploration Research Group Phase I & II Team;
RA "Analysis of the mouse transcription based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002);
DR EMBL; AK088653; BAC04080.1; -.
FT NON_CODING
FT SEQUENCE 799 AA; 88814 MW; 415D3E687150A59A CRC64;

Query Match      36.6%; Score 1178.5; DB 11; Length 799;
Best Local Similarity 48.1%; Pred. NO. 7.2e-97;
Matches 255; Conservative 74; Mismatches 182; Indels 19; Gaps 11,

OY 66 YVGAREALFAF-SMALEILOGAISAEVAVERKTECIOGKNNOFCFPFIPLPYNAH 124
   |||||...::|...:||||:||||:||||:||||:||||:||||:||||:||||:
DB 1 YVGAENAFVAVNALNVEYKHIEHYKVSEEDKSKCAEKSGKSQCETCLNYIRVDLPISRTS 60
   |||||...::|...:||||:||||:||||:||||:||||:||||:||||:||||:
125 LVVCCTVAFPCPKTYVMNLTFTLHEGEFEEDGGCKCPYPDAKGAGLLVDGLYSATLNKF 184
   |||||...::|...:||||:||||:||||:||||:||||:||||:||||:||||:

```

```
DB 61 LVYCTNAFOTPCDHLNLSRFF-LGKSEDOCKGRCPFPBANSYSTVVMGGLYSTSYNF 119
OY 165 LQTEPIILRNKQPHNSMTEYLAFWLNBPFGVSAVYVESVSGFTGDDKYFFERERAY 244
    ||||| ||| :||| ||||| :||| :||| ||||| :|||
DB 120 LGSEPIISRN-SHSPLRTEREYAPWLNBPFGVSAVYVESVSGFTGDDKYFFERERAY 178
OY 245 ESDCAEBOVAVARVAVCGDNGARTLOKRTTFKAKALACAPRMWOLYNLOLQAMHILQ 304
    :||| :||| ||||| ||| ||||| ||||| ||| :||| :|||
DB 179 EYEFVFKLMPVAVAVCGDNGARTLOKRTTFKAKALACAPRMWOLYNLOLQAMHILQ 238
OY 305 DTSWNTFFGVFOQMGDMYLSAICEYOLEIQRNF-EGPYKE--THEEAKQDRYTD 360
    :||| :||| ||||| ||| :||| :||| ||||| ||| :|||
DB 239 APGLKEPFAVAVFTQOLNNGVSAVCAATLAEVAFSRGKMGATQVQSITKVRING 298
OY 361 PVPSPRSCJNMNHRHGYTSSLEPDIINLFEVKKHPLMEQVPRMSRPLLVKGTNF 420
    ||||| ||||| :||| ||||| ||||| ||||| ||||| |||||
DB 289 PVPTRPACIDISEARAAVNTSSNLDPKTLQFVNDHPLMDSVPRIDNRPLTKKDVY 358
OY 421 THLVADRTGLDGAATYTLFTGTGDMYLLKAVSLQPMVHLTELDLF-DQEPMSLVLSQ 479
    :||| :||| ||||| ||| ||||| ||||| ||||| |||||
DB 359 TQIVVDROALDGTGYDVFISTDRGALHKAVALKEVHVEITGLFRDSEPVTLILSS 418
OY 480 SK-KLIFAGSRSLVQLPVADCIKYRSCADCVLARDPYCAMSNTSRCAVAGHFGSL- 536
    :||| :||| ||||| ||| ||||| ||||| ||||| |||||
DB 419 KKGKRFVAVAGSVVQAPLAFCEKHGSCDQCVLARDPYCAMSNTSRCAVAGHFGSL- 478
OY 537 -LIOHMTSDTSGICNLGSKIGSPKRNITVAGTDLVLPCHLSNLA 585
DB 479 GWIOD-MSGDITSS-CLDKSEFNOHFKEH-----GTAELKCFQKSNLA 521
```

RESULT 6

```
DB ID Q8BIC3 PRELIMINARY: PRT: 550 AA.
AC Q8BIC3;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Semaphorin 4B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK028896; BAC26181.1;
SQ SEQUENCE 550 AA; 61309 MW; 8C28E29DCAAC958 CRC64;
```

```
Query Match 36.3%; Score 1169.5; DB 11; Length 550;
Best Local Similarity 46.2%; Pred. No. 2.7e-96;
Matches 239; Conservative 87; Mismatches 172; Indels 19; Gaps 8;

OY 27 LVPRKTVSSGELATVVRPSOTGIODFLTLTEPTGLVYGAREALFAFSEALELGA 86
    :||| :||| ||||| ||| :||| :||| ||||| ||| :|||
DB 31 LQPRISVPLSGEERLIRKEFAENISNTYALLISQCKTLVYGAREALFALNSFLPG 90
OY 87 ---ISMVAPEKTECIQKGNKNOTECFNFIRLOPYNASHLYVCGTYAFQPKCTVYM 142
    :||| :||| ||||| ||| ||||| ||||| ||| :||| :|||
DB 91 EYQELMSADNRKQCKSPKGDPRKDCQNYIKILPLNSHLLVCGTAASFPLCATYHI 150
OY 143 LFTLLEHGE----FEDGKGCYPDPKAGHAGLVYDGLYSATLNFQTEPIILRNKGP 197
    :||| :||| ||||| ||| ||||| ||||| ||| :||| :|||
DB 151 ASFTLADQEGAVNLEDEKGRCPDPNFKSTALVYDGLYGTGVSSFOGNDPAISRQSS 210
OY 198 HHSMTETLAFWLNBPFGVSAVYVESVSGFTGDDKYFFERERERAYSDCAEBOVAV 257
    :||| :||| ||||| ||| ||||| ||||| ||| :||| :|||
```

```
DB 211 RPT-KTESSLNKLQDPAPVAVSAVYVESVSGFTGDDKYFFERERENTIVSRV 269
OY 258 AVACGDMGARTLOKRTTFKAKALACAPRMWOLYNLOLQAMHILQDT--SWNTTFEG 315
    ||||| ||| :||| ||||| :||| :||| ||||| :|||
DB 270 AVACGDEGEVAVLQRMNLSFKADQLSRPDGEPFNVLQDVLPNRPDQMRKTLFYG 329
OY 316 VFOAOV--GDMYLSAICEYOLEIQRNFEGPYKEHEEAKQDRYTDVPSRPSQINN 373
    ||| :||| ||||| ||| ||||| ||||| ||| :||| :|||
DB 330 VFTSQMHRRTGSAICVFTMNDVOKAFGLKKKVRRETQWYTFHOVTPRPACITN 389
OY 374 MRRRGYTSLEPDIINLFEVKKHPLMEQVPRMSRPLLVKGTNFTHLVADRTGLDG 433
    :||| :||| ||||| ||| ||||| ||||| ||||| |||||
DB 390 SAREKRNSSLDLPQVAVLNFQDHLMDQVR---SRLLDLPRAQVAVHVRPGLH- 445
OY 434 ATTYVLFTGTGDMYLLKAVSLQPMVHLTELDLFDQ-EPMSRVLVSQSKLLFAGSRQL 492
    :||| :||| ||||| ||| ||||| ||||| ||||| |||||
DB 446 STYDVLFTGTGDMYLLKAVSLQPMVHLTELDLFDQ-EPMSRVLVSQSKLLFAGSRQL 505
OY 493 VOLPVADCIKYRSCADCVLARDPYCAMSNTSRCAVAV 529
DB 506 VQVPVAVNSLYPTGCDCLLARDPYCAMSNTSRCAVAV 542
```

RESULT 7

```
DB ID Q8WU9 PRELIMINARY: PRT: 761 AA.
AC Q8WU9;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Colon;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020974; AAH20974.1;
DR InterPro: IPR000005; HTHARC.
DR InterPro: IPR002165; Plexin-like.
DR InterPro: IPR001627; Sema.
DR Pfam: PF01437; PSI; 1.
DR Pfam: PF01403; Sema; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
KW Hypothetical protein.
SQ SEQUENCE 761 AA; 83573 MW; 5BCB889AA32A2B3 CRC64;
```

```
Query Match 32.0%; Score 1029.5; DB 4; Length 761;
Best Local Similarity 40.2%; Pred. No. 1.7e-83;
Matches 237; Conservative 69; Mismatches 217; Indels 47; Gaps 15;

OY 28 VPRKTVSSGELATVVRPSOTGIODFLTLTEPTGLVYGAREALFAFSE---ALETQ 84
    :||| :||| ||||| ||| :||| :||| ||||| ||| :|||
DB 39 MRRVAVYADDERALSFHQKLODPEDTLILSGDNTLVYGARALALDIDDPVPRK 98
OY 85 GAISDAPEKTECIQKGNKNOTECFNFIRLOPYNASHLYVCGTYAFQPKCTVYM-- 142
    :||| :||| ||||| ||| ||||| ||||| ||| :||| :|||
DB 99 NMIPAPASDRKSECAFKKSNETQCFNIRVLVSYNVHTLTCGTFARSPACTFELQD 158
OY 143 -LFTLLEHGEFBDKGCYDPKAGHAGLVYDGLYSATLNFQTEPIILRNKGRHSM 201
DB 159 SYLLPISDQKVEGQSPDPNFKSTALVYDGLYGTGVSSFOGNDPAISRQSS 218
OY 202 KTEYLAFWLNBPFGVSAVYVESVSGFTGDDKYFFERERERAYSDCAEBOVAVARV 260
    :||| :||| ||||| ||| ||||| ||||| ||| :||| :|||
DB 219 KTDNRLRLHMDASV--AAIST-----QVYTFEEETASQDFPERLHTSRVAV 268
```

```

OY 261 CKDGMGARTLQKRTTFLKARLACGAPNQLTFNQLQAMHTLQDTSMHNTFFGCFQAO 320
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 269 CKNDVGEKILQKRTTFLKARLACGAPNQLTFNQLQAMHTLQDTSMHNTFFGCFQAO 327
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 321 W--GDWMLSAICVQLEIQRVEGPGPKYEEHEAOKMDRTDPPSPRSCINMRRH 378
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 328 MGVGGRSSAVCAFSLDIEVEKGYKELNKTSSRTTYRGPTNDRPGSC----- 379
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 379 GYTSLELPNILNFKHPMEBOVGRMSRPILVKKGTFTLHVDRLVADGAYTV 438
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 380 ---SVGSSDKALFMDFHLMDEQ---VGPLLVASGEYTLAETAGLDGSHLV 433
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 439 LFTGTGGMILKAVSLG--PWHVLEELQF--DOEPMSLVLSQSKLLFAGSRLVQLP 496
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 434 MYLGTGTSLSHKAVSSDSSAHLEVEIQLFPPDPPEVRLQLAPGAVFVFGSGVWVP 493
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 497 VADICRSCADCVIADPYCAMSVNTSRCAVAGHGSLLIOHVMHSDTSGIC----NL 552
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 494 RANCVSYESCVDYADLPDPCAMPDPSRTCCLLSAPVLSKWKODMERGNPEMACASGPM 553
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 553 RGSRTIOGCP--XPKNITVAGTDLVLP--HLSSNLAL-----PDSNPRESS 595
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 554 RSLRPGSRPQILKEVLAVPNSITELPCPHLSALASYWMSHCPAVPEASS 603
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 8
O8BTW9 PRELIMINARY; PRT: 774 AA.
ID 08BTW9
AC 08BTW9
DT 01-OCT-2002 (TREMblrel. 22, Created)
DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Semaphorin 3A.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Head.
RA Tannahill D., Nielsen J., Regan A.G.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY030051; AAK8166.1;
DR InterPro: IPR003559; Ig.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001005; MyD_DNA_binding.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR001627; Sema.
DR Pfam: PF00047; Ig_1.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00409; Ig; 1.
DR SMART: SM00423; PSI; 1.
DR SMART: SM00630; Sema; 1.
DR PROSITE: PS00835; IG_LIKE; 1.
DR PROSITE: PS00037; MYB_1; 1.
SO SEQUENCE 774 AA; 89319 MW; 3127ED2E0CCD5FA4 CRC64;

Query Match 28.4%; Score 913; DB 13; Length 774;
Best Local Similarity 39.3%; Pred. No. 5.2e-73;
Matches 204; Conservative 88; Mismatches 195; Indels 32; Gaps 13;

```

```

DB 147 GHVEODNVKLEESFFENGKSPYDPKLLTASILIDGELYSTADFGRFARFRTL 206
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 196 GPHSMKTE--YLAFLNPHVGSAYVPEVSGSTGDDVYFFERRAAYSDCYAQV 254
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 207 GNNHPIRTEODHSDMLNDPRIGAYLVPSNP--BDDVYFFERRALDGEHTGATH 263
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 255 ARVAVCKGDMGARTLQKRTTFLKARLACGAPNQLTFNQLQAMHTLQDTSMHNT 311
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 264 ARIQOLCKNDPGHRLVNNKTFLLKARLCSVGPBGIDPHDELVDVILNSKDRNP 323
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 312 TFGVFOAQMDMTLSAICVQLEIQRVEGPGPKYEEHEAOKMDRTDPPSPRSC 371
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 324 VYVAVFTSSNIFGSAVCMSTLSDIRVFLGPAHBDGPYQVPGQVPRPDRPGTGP 383
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 372 NNNHRRHGYTSLELPNILNFKHPMEBOVGRMSRPILVKKGTFTLHVDRLVAD 429
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 384 RQTF--GQFDSKLPDDEVIFARSHRAMYPVPINNRPIITTEVDYQFTQVVDRE 441
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 430 GLDGATYVLEFGDGMILKAVSL--GPVH---LIEELQDOEP--NRSVLSSQK 481
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 442 AEDG--QYDVMETIGDMGILKAVSVPKETWTDLEVLLEMTVF--REPTAISAMEIS 499
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 482 KLFAGRSQVLQVLPVADICITV--RSCADCVIADPYCAM 519
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 500 QOLYIGSSVGSQPLPHRCDVYGRACACCIADRPYCAM 538
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 9
O8TB71 PRELIMINARY; PRT: 748 AA.
ID 08TB71
AC 08TB71
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC024220; AAB24220.1;
DR InterPro: IPR003599; Ig.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR001627; Sema.
DR Pfam: PF00047; Ig_1.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00409; Ig; 1.
DR SMART: SM00423; PSI; 1.
DR SMART: SM00630; Sema; 1.
DR PROSITE: PS00835; IG_LIKE; 1.
DR Hypothetical protein.
SO SEQUENCE 748 AA; 83034 MW; 587C53CB5A4656 CRC64;

Query Match 27.5%; Score 884.5; DB 4; Length 748;
Best Local Similarity 36.1%; Pred. No. 1.8e-70;
Matches 214; Conservative 92; Mismatches 24; Indels 39; Gaps 19;

```

```

Db 132 GTGAHPTCAFEVGHRAEPEVRLDPGRIGEDGKGSPPDPRIHRAASVLYGELISGVAA 191
183 NFIAGTEPILIRNMGRPHSMKTE-YLAFLMNEPHEVGSAYVPESVSGFTGDDDKVYEFNE 241
192 DLAMGNDFTIFRLGQRPRLRTERPHDSRWMLNEKFKYAVFWIPESNP---DDDKIYEFNE 248
OY 242 RAVE-SDCAEOVAVANAVKCGMDMGATLQKRWTFELKARLACSAAPMWQ--LYENQLO 298
Db 249 TAVEAPALGRLSVSVSGQICRNVDGQSLVKNKMTTFLKARLYCSVPVEGDTHTDQLO 308
OY 299 AHHITQDTSVHNTTFEGVQAQMGDMYLSAICEYOLEEYORVEEGYKEXHEAOKWXY 358
Db 309 DVFLLSSRDHRTPLLYAFVSTIS-STEFGSAVVCYSNNDVRAFLGFAKREGPMQWASY 367
OY 359 TDVPSPRGSCINNMHRRHG-YTSLELPELNLNFKPLMEBOVGRPMSPVLYKRG 417
Db 368 QGRVYPRGKPS---KTFGTFSSTKDPDPDVIQPARNHPKMYNSVLTGGRPLEYVG 424
OY 418 TN--FTHLVADRVTGLGATYTVVLEFGDGMLLANVSLGPMVH-----LIEELQLF-D 468
Db 425 ANVTFTQIADBVAAADG-HYDVLFTGTGVTLKAVISVPGKSRPSABGLLLEELHVFED 483
OY 469 QEPNRSVLSQSKLLFAGSRSLQVLPVADCIK-RSCADCVLARDPYCAN-SVNTSRC 526
Db 484 SAAVTSKMOISSKRHQLYASBSAVAOIALHRCAGRGVCTECCCLADPYCAMDGVACTRF 543
OY 527 VAVGHPFSLIQHVTSDTSGICMLRSGKISGHPKNIYVAGTDLVPC 578
Db 544 QPSAKR--RFRQDVANDPSTLCS--GDSSRPALLEHNVFVEGSSATLEC 591
RESULT 10
OQBH34
ID QQBH34 PRELIMINARY; PRT; 777 AA.
AC 08BH34:
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Semaphorin 3D precursor homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid-10090;
RN 11)
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J; TISSUE-Kidney, and Skin;
RX MEDLINE-22354683; PubMed-12466851;
RA The FANTOM Consortium.
RT the Riken Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RT Nature 420:563-573(2002).
DR EMBL; AK028900; BAC26185.1;
DR EMBL; AK028900; BAC26185.1;
DR EMBL; AK028900; BAC26185.1;
SQ SEQUENCE 777 AA; 89548 MW; 5450DBD45D1DBAF CRC64;
Query Match 27.3%; Score 878.5; DB 11; Length 777;
Best Local Similarity 38.6%; Pred. No. 6,6e-70;
Matches 193; Conservative 88; Mismatches 190; Indels 29; Gaps 14;
OY 46 SQTGQDVLTLTLPRTGLLYGAREALFAFSEMLELD-QAISEAPVEKTECIQK 104
Db 64 SSEGL-DEQTLILDERGILLGANDHVELLSLVLNKFKKIYPAKEREVELCKLKG 122
OY 105 NNQGTCFNFIPLQYNSHLVYCGTVAFOPKCTVNM-----LFTLEHGEFEDGK 158
Db 133 DANACAFPIKLYQDYNTHTYVCGTGAFFPLCGIITDGAKEELIFKLDTHNLSGRK 182
OY 159 CPYDPAKAGLVLVDELVSATLNNFLCTEPIILRNMG---PHNSKTEYLA-FWLNPH 214
Db 183 CPFDQGFASVMTDEHLYSGTASDFLGDKTAFTSTSLGMDHHSIRDISHHNLNKA 242

```

```

OY 215 FVGSATVPESVSGFTGDDDKVYEFREAVESDCAEAOVAVARVCKGDMGARTLQK 274
Db 243 FICTFPIPD---TYNPDODKIFYEFFRESSQEGSTSDRSLTSRGVRCNDVQGRSLNK 299
OY 275 WTFELKARLACSAAP---NMOLYNOQAMHTLQDTSVHNTTFEGVQAQMGDMYLSAICE 331
Db 300 WTFELKARLACSAAP---NMOLYNOQAMHTLQDTSVHNTTFEGVQAQMGDMYLSAICE 331
OY 332 YOLEEIORVEEGYKEXHEAOKWXY-TEFLVADRVTGLGATYTVVLEFGDGMLL 449
Db 360 YSMADIRAVFNGYFAIKESDHRMVOYDGRIPFRGTCTPSKTYDL-ISTDFPDVY 418
OY 392 NFYKHPLEEDVGRPMSPVLYKRGTF--TEFLVADRVTGLGATYTVVLEFGDGMLL 449
Db 419 SFTRRPVYKSYVPAAGAPTFKRI NVDYRLTOIVVDHYVAEDG-QYDVAFGLTDIGTVL 477
OY 450 KAVSLG-FW---VHLEELQLFQ-EPKRSVLSQSKLLFAGSRSLQVLPVADCIK 503
Db 478 KAVSISKERNNEEVLLEELQVFKHPTAILMELSLKQOLYVGSMDGLVQSLHRCDTY 537
OY 504 -RSCADCVLARDPYCAMSVN 522
Db 538 GKACADCCCLADPYCAMDGN 557
RESULT 11
OQBTV7
ID OQBTV7 PRELIMINARY; PRT; 754 AA.
AC 08BTV7:
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Semaphorin 3B.
GN SEMA 3B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid-9606;
RN 11)
RP SEQUENCE FROM N.A.
RA Koyama N.;
RT "Semaphorin 3B (SEMA3B) cDNA."
RT Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB083186; BAB88870.1;
DR InterPro: IPR003599; Ig.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig-MHC.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR01627; Sema.
DR Pfam: PF00047; Ig. 1.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00409; Ig. 1.
DR SMART: SM00423; PSI. 1.
DR SMART: SM00630; Sema; 1.
DR PROSITE: PSS0835; IG_LIKE. 1.
SQ SEQUENCE 754 AA; 83691 MW; E61FD08C04E9A68E CRC64;
Query Match 27.2%; Score 876.5; DB 4; Length 754;
Best Local Similarity 35.7%; Pred. No. 9,5e-70;
Matches 213; Conservative 93; Mismatches 248; Indels 43; Gaps 19;
OY 12 ARKGLGIGAEYWMNLVPRKTVSSGELAT--VVRFSQGTGIDFTLTLTLPRTGLLYGA 69
Db 14 ALLNMGVGLSAA--PSPRLRLSFQELQAMHGLQFSLRFTCCYQALLVDEERGRLVGA 71
OY 70 REALAFSEMLELDG-AISMEAPVEKTECIQKNNOTCECFNFIPLQYNSHLVYCG 128
Db 72 ENHVASLNDNISKRAKLLAMPAPVEMREECMAGADITCEKNPVKLLHAARHTLLAC 131
OY 129 GTYAFOPKCTVNM-----LFTLEHGEFEGDQKGCSPYDPAKAGLVLVDELVSATLN 182
Db 132 GTGAHPTCAFEVGHRAEPEVRLDPGRIGEDGKGSPPDPRIHRAASVLYGELISGVAA 191

```

```

OY      183  NELGEPILINNMGGHNSKTE-YLAFMLNETHFUGSAVPEVSFTGDDGXVEFFRE 24.1
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      192  DLMGDFITFSLGQRPILKTEPHDSRMLNEKFAVMIPESEN---DDDKITFFFRE 24.8
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      242  RAVE-SDCYAEQVAVARAVKCKMGDGGATTLOKMTTEFLKARLCAPMWO--LYFNOL- 29.7
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      249  TAVEAAPALGRLSVRSVQIGKANDVGQSLVNKKMTTEFLKARLVCSVPQESGTHFDLR 30.8
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      298  ----QAMHTLODTSHNHTTFEGVCPQAQCKMTLSAICEYQLEIQVFEBSPIKYEYHEAO 35.3
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      309  PEPAPADVLLSSRDHRTPLLYAVFSTSSSIFQGSACVYSMDVBRALFGLPEAHKCKGPMH 36.8
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      354  KMDRTDVPSPRPGSCINNMHRRNG-VYSSLELPIINLIFYKKHMLMEQVCPKMSRPL 41.2
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      369  OMVSTQGRVPIPRGMCPS---KTEGTSESTLDPDPDVIQFANHPPLMNSVLTPTGRRPL 42.5
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      413  LKKKGTN--ETHLVADRYVGLDGAFTYVLEIGTDGMLLKAVSLGPMVH-----LIEEL 46.4
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      426  FLOWGANTFTFIADRYVAADG-HYDVLFIGIDVETVLKVISVPGKSRSAAGLILLEL 48.4
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      465  QLF--DQEPKRSIVLSQSKKLFEPAGRSQVLQLPVADCIKY-RSCADCVIARDPYCAM-SV 52.1
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      485  HFVEEDSAVTSQOISSKRQQLVYASRSAAVOIALHRCAMHGRVCTECCIARDEYCAMDGV 54.4
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      522  NTSRCVAAAGGHGSLILQHVMNTPDSGICNLNGSKIQSGPXPKNITVVAAGTDLVLR 57.8
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      545  ACTRQPSAKR--RFRRODVNRNGDPSLTCS--GDSRPALELHKVGVRESSNAFLEC 59.7
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 12

ID Q8BMF6 PRELIMINARY; PRT; 777 AA.

DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Semaphorin 3D precursor homolog.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 ON NCBI_Taxid=10090;
 RN [1]
 RP
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK031704; BAC27522.1; -
 SO SEQUENCE 777 AA; 89562 MW; B7204DB8228B89CE CRC64;

Query Match	27.1%	Score 872.5;	DB 11;	Length 777;
Best Local Similarity	38.4%;	Pred. No. 2.3e-69;		
Matches 192; Conservative	88;	Mismatches 191;	Indels 29;	Gaps 14

```

OY 46 :OTG:ODELT:LTLETPF:TLVYGA:EAELFA:MSMEAL:ELQ-GA:ISNEA:VEKTE:TEL:GKX 104
Db 64 :SSEGL-DY:OTLL:DEER:GILL:GA:NDHVF:LS:VDLNNFK:KITYPA:KAREL:ELCKL:AG 122
OY 105 :NNGEC:ENF:IRFL:OP:YNA:SHLYVC:GTVA:OP:KCTV:YNN-----LT:FL:HEHEF:EDG:GK 158
Db 123 :DANAC:ANF:IRV:QY:NKTHVY:CGTGA:PH:LCGY:ID:GANK:EE:LF:MD:PH:NL:ES:GR:K 182
OY 159 :CPYPAK:GHAG:ILVD:GELX:SATL:NFL:TEP:IL:NNG---PH:SKTE:YTA-LFWL:NEPH 214
Db 183 :CPFD:QO:CFAS:MTDE:HLV:SGTAS:PF:LRDTA:FT:SLG:SL:MD:HN:SI:RFD:ISEH:WLN:GAK 242
OY 215 :FVGS:AV:SVES:GFTG:DDDK:YFF:FE:RE:RA:VES:DCVA:EO:YVA:VVA:VAV:G:GDG:MA:RTL:QK 274
Db 243 :F:IG:FP:DPD---T:Y:PD:DK:IT:FE:RES:Q:ES:TS:DR:SL:SV:GVC:KND:VG:GQ:SL:LNK 299

```

[illegible]

RESULT 13

ID	Q9QX23	PRELIMINARY;	PRT;	775	AA.
----	--------	--------------	------	-----	-----

DT 01-MAY-2000 (TREMBLERL. 13, Created)
DT 01-MAY-2000 (TREMBLERL. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLERL. 23, Last annotation update)
DE Semaphorin M-Semak.
GN SEMA3E.
OS Mus musculus (Mouse).
OC Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57/Black 6;
RA Miyazaki N., Furuyama T., Inagaki S.;
RT "A novel semaphorin, M-Semak which inhibits neural outgrowth from
RT sensory neurons."
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF034744; AAD01996.1; -.
DR MGD: MGI:1340034; Semase.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR007110; Ig-1like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR001627; Sema.
DR Pfam: PF00047; Ig. 1.
DR Pfam: PF01403; Sema. 1.
DR SMART: SM00409; Ig. 1.
DR SMART: SM00423; PSI. 1.
DR SMART: SM00630; Sema. 1.
DR PROSITE: PS50835; IG_LITE. 1.
SQ SEQUENCE 775 AA: 89543 MW: 221E766FA0409BD4 CRC64;

Query Match	26.58;	Score 852.5;	DB 11;	Length 775;
-------------	--------	--------------	--------	-------------

Matches 210; Conservative 86; Mismatches 216; Indels 61; Gaps 22;

[illegible]

```

Db 176 SSVSTLGNELFAGLSDYWGGRDIAFRSMGKLCHIRTEHDERL---LKEPKFVGSTW 232
OY 221 VPEVSGSTGDDDKYFFEFERAVESDCYAEQVAVARAVCKGDMGARTLOKRTTFLK 280
Db 223 IPDNDR---DNNKMYFFETEKALAEENNAHTIYRVGLICVNDMGGRILVKNKSTFLK 289
OY 261 ARLACSAAPNW---QLFYNOLOAMHTLODTSWHTTFFGVFOAQOMDMYLSAICEYOLEBT 337
Db 290 ARLVSVGMGIDYFDELEDEYFLLPTRDPKNPIYFGLFNTTSTNIFRGHACVYHMSI 349
OY 338 QVFECPY-----KEYHEAOKMDRYTDPVSPRPGSCINNMHRRHG--YTSLELPINI 390
Db 350 REAFNGPFAHKEGPRYH---WSLYEGKVPYPRPGSCAS---KVGNGKYGTTKDYPPDA 401
OY 391 LNFVKNHPLMEQVGRPMRSRLPVKKG--TNEFHLVADRYTGLDGTATYVLEFGTGDWL 448
Db 402 IRFAMHPLMWQPIKPYKPKPIVKTGDKYNLRQLAVRVEADG-QYDVLVLTGDTGIV 460
OY 449 LKAVSL---GPVNH-LIEELQLF-DQEPMSLVLSOSKLLFAGRSOLVOLFVADCI 501
Db 461 LKVIITYNOETEMMEVILEELQIFKDPAPITSEISSKROOLYIGSASAVAVQVRFHCD 520
OY 502 KYRS-CADCVLARDPYCAM-SVNTSRGVAVGSH 532
Db 521 MYSACADCCCLARDPYCAMDGTICSRYPYTGAN 553

RESULT 14
O9NS98 PRELIMINARY: PRT: 782 AA.
AC O9NS98; Q9H7Q3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Semaphorin sem2 (FlvJ00014 protein).
GN SEM2 OR FLVJ00014.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Plimates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Seki N., Hattori A., Hayashi A., Kozuma S., Muramatsu M., Miyajima N.,
RA Salto T.;
RT "Human semaphorin.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 58-782 FROM N.A.
RC TISSUE=Spleen.
RA Ohara O., Nagase T., Kikuno R., Okumura K.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
RT spleen.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029496; BAA98132.1; -
DR EMBL; AK024425; BAB15715.1; -
DR InterPro: IPR003599; IG.
DR InterPro: IPR007110; IG-1like.
DR InterPro: IPR003006; IG-MHC.
DR InterPro: IPR003659; Plexin-1like.
DR InterPro: IPR001627; Sema.
DR Pfam: PF00047; Ig_1.
DR Pfam: PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
SQ SEQUENCE 782 AA; 86700 MW; 85CB424874DF6663 CRC64;

```

Query Match 26.5%; Score 851.5; DB 4; Length 782;
 Best Local Similarity 36.6%; Pred. No. 1.8e-67;
 Matches 204; Conservative 89; Mismatches 210; Indels 55; Gaps 16;

```

OY 1 MAPH-WAV-WLLAARLKGIGCAEYWMNLVPRKTVSSGELATVVRFSQTGIQDFLTTL- 57
Db 1 MAPSAMALICWLLGGLLIGSGSSGSPGSPVPLRLSLRLLSANSATFLGPGSLNQA 60
OY 58 --LTPEPTGLLYGAREALFAFMSALELQALSWENAPV-----KTECIQGNKNG 107
Db 61 MYLDEYRDRFLGGLDIALSLRLDQ-----AMPDPREVLMPPOPGQRECEVNRGRDPL 113
OY 108 TPCFNFIRELPQYNASHLYVCGTYVAFOPKCTYVNM-----LTFTLEHGEFEDGKRCPYD 162
Db 114 TGCANFVRLQPHNTHLLAGCTGAFQPCALITVGHREHNLHLEPGSVESGRRCRPE 173
OY 163 PAKGAGLLVDCELYSATLNNFLGTEPIILRNMGPHSKTEYLAFWLNEHPVGSAYVP 222
Db 174 PSRPASPTFIDELTYGLADFLGREAAMIFRSGGRPALRSDQSLLDHPRTFYAAARIP 233
OY 223 EIVSGFTGDDDKYFFEFERAVESDCYAEQV-VAVARVCKGDMGARTLOKRTTFLK 281
Db 234 ENSDO---DNDVYFFSETVPSPDGSHVTVSRGVCVNDAGGQRLVKNKSTFLKA 290
OY 282 RLACGAP---NMQLYFNOLQAMHTLODTSWHTTFFGVFOAQOMDMYLSAICEYOLEEIO 338
Db 291 RLVCVSPGPGALFTEHDOLEDEYFLLMPKAGSLLEYALFSTVSAVFGGFAVCYVHADIW 350
OY 339 RVFEQPKYEHAEQAKMDRYTDPVSPRPGSCINNMHRRHG--YTSLELPINIIPYAK 396
Db 351 EVFNGPFAHKEGPRYH---WSLYEGKVPYPRPGSCAS---KVGNGKYGTTKDYPPDA 410
OY 397 HPLMEQVGRPMRSRLPVKKG--TNEFHLVADRYTGLDGTATYVLEFGTGDWL 449
Db 411 HPLMFPVPRHGRVLRK-----THLAQQLQIYVDRVEADG--TYDVFLEGTSGSLV 464
OY 450 KAVSL---GPVNH-LIEELQLF-DQEPMSLVLSOSKLLFAGRSOLVOLFVADCI 502
Db 465 KYVALQAGSAPREEVLEELQVFKVPYPTTEMEISVKNQMLYVSGRLGVALRLHQCT 524
OY 503 Y-RSCADCVLARDPYCAM 519
Db 525 YGTACAECCLARDPYCAM 542

RESULT 15
O8OGU9 PRELIMINARY: PRT: 756 AA.
AC O8OGU9;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Semaphorin 3f.
GN SEMA3F.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe Y., Nakamura H.;
RT "Axon guidance of the trochlear nerve by Sema 3f along mid-hindbrain
RT boundary.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB072930; BAB88691.1; -
DR InterPro: IPR007110; IG-1like.
DR InterPro: IPR003598; IG-C2.
DR InterPro: IPR003006; IG-MHC.
DR InterPro: IPR003659; Plexin-1like.
DR InterPro: IPR001627; Sema.
DR Pfam: PF00047; Ig_1.
DR Pfam: PF01403; Sema; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS50835; IG_LIKE; 1.

```


Immunoglobulin domain.
SEQUENCE 756 AA: 85196 MW: C36754C02541ED88 CRC64:

Query Match 25.4%: Score 818; DB 13: Length 756;

Best Local Similarity 35.5%: Pred. No. 1.7e-64;
Matches 195; Conservative 86; Mismatches 217; Indels 52; Gaps 16;

```
QY 6 AVWLLAARLWGLGIGAEVWNL-----VPRKTVSSGEL-ATVVRFSQ--TGIDF 53
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 4 ATVLLMTLTLTG-----WRAHGKDGVPPTPRVOLSFELKATGTAHFNFPLNSDY 57
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 54 LTLTLTEPTGLLYGAREALFAFSEALDELQ-ALSWEPAYEKTECIQKKNQTECFN 112
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 58 RILKEDHDHMYGSKDYLSLDLHDINREPLIHPASQORIEECILSGKNSNGECGN 117
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 113 FIFELQPNASHLYVCGTYAFOPKCTVYNM-----LFTLEHGFEFGKGCPCYPDAKG 166
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 118 FIFLQPNRTHLYVCGTGATNPICAFINRGRKADYIFYLEPKLESKGKCSYDPKVD 177
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 167 HAGLLVDGELYSATLNNFLGTEPIILRNMGPHSMKT-EYLAFWLNEBHFVGSAYVPEV 225
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 178 TVSALINBELYAGYIDFMGTDAIFRTMGKOTAMRTQYNSRWLNDPAFVRAQLPD-- 235
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 226 GSFTGDDDKYFFFRRAVESDCYAEQVAVARVAVCKGDMGARTLQKRTTFELKARLAC 285
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 236 -SSERNDDKLYFFFRKSADAP-LSPGYSRIGRICLNDGHCCLVNMKSTFLKARLYC 293
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 286 SAP---NMQLYFNQLQAMHTLQDTSMHNTFFGVFOAOWGDMYLSAICEYQLEFIQVFE 342
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 294 SVFGRPDGIETHEDELQVFIQOTDTKNPVIYAFVSAGSVFKSAYCVYSMADIRNVFN 353
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 343 GPYKEYEAEAKMDRYTDPVPSPPRGSCINNMRHRYTSSL---ELPDNILFVKKRP 398
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 354 RPFHKKEGPNQOMPYGKMPYRPGTQPG-----TFTPSMKSTKDYDPDEVINMRSH 408
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 399 LMEQVGRPMGRPLLYKKGTN--FTHLVADRYTGLDQATYVLEIGTGDGHLKAVSLGP 456
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 409 LMYHAVYPAHRQPLVFRTNVYRFTTIAVDQVDAAG-RYEVLFGTDRTGTVQKVIYLP 467
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 457 WVHLIEELQLFDOE-----PMRSILVLSQSKLLFAGRSQVLOVPVADCIKY-RSCADC 509
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 468 DDMETEELMLEIEVFVPAPIKMTISSKQOLYSSAVGVTHLALHRCDYGEACADC 527
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 510 VLARDPYCAW 519
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 528 CIARDPYCAW 537
```

Search completed: August 7, 2003, 13:49:56
Job time : 109 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2003, 13:46:05 ; Search time 30 Seconds

(Without alignments)
840,576 Million cell updates/sec

Title: US-10-002-050-14

Perfect score: 3219

Sequence: 1 MAFHMAVILLARLKGIG.....PCHLSNIALPDNSPESSV 596

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1220.5	37.9	862	4	US-08-556-422A-2
2	1149	35.7	607	4	US-08-556-422A-4
3	932.5	29.0	655	4	US-08-556-422A-3
4	932.5	29.0	771	1	US-08-121-713D-54
5	932.5	29.0	771	1	US-08-835-268-54
6	932.5	29.0	771	2	US-09-060-692-54
7	932.5	29.0	771	3	US-08-833-391-54
8	932.5	29.0	771	4	US-09-060-610-54
9	932.5	29.0	771	5	PCT-US94-10151A-54
10	890.5	27.7	425	4	US-08-556-422A-7
11	852.5	26.5	775	4	US-09-308-179B-1
12	727.5	22.6	930	4	US-09-254-594-6
13	719.5	22.4	295	4	US-08-556-422A-6
14	704.5	21.9	730	1	US-08-121-713D-58
15	704.5	21.9	730	1	US-08-835-268-58
16	704.5	21.9	730	2	US-09-060-692-58
17	704.5	21.9	730	3	US-08-833-391-58
18	704.5	21.9	730	4	US-09-060-610-58
19	704.5	21.9	730	5	PCT-US94-10151A-58
20	685	21.3	724	1	US-08-121-713D-62
21	685	21.3	724	1	US-08-835-268-62
22	685	21.3	724	2	US-09-060-692-62
23	685	21.3	724	3	US-08-833-391-62
24	685	21.3	724	4	US-09-060-610-62
25	685	21.3	724	5	PCT-US94-10151A-62
26	680	21.1	929	4	US-09-254-594-3
27	658	20.4	650	1	US-08-121-713D-60

28	658	20.4	650	1	US-08-835-268-60	Sequence 60, Appl
29	658	20.4	650	2	US-09-060-692-60	Sequence 60, Appl
30	658	20.4	650	3	US-08-833-391-60	Sequence 60, Appl
31	658	20.4	650	4	US-09-060-610-60	Sequence 60, Appl
32	658	20.4	650	5	PCT-US94-10151A-60	Sequence 60, Appl
33	654.5	20.3	887	4	US-09-077-940A-2	Sequence 2, Appl
34	646	20.1	888	4	US-09-077-940A-4	Sequence 4, Appl
35	596.5	18.5	712	1	US-08-121-713D-64	Sequence 64, Appl
36	596.5	18.5	712	1	US-08-835-268-64	Sequence 64, Appl
37	596.5	18.5	712	2	US-09-060-692-64	Sequence 64, Appl
38	596.5	18.5	712	3	US-08-833-391-64	Sequence 64, Appl
39	596.5	18.5	712	4	US-09-060-610-64	Sequence 64, Appl
40	596.5	18.5	712	5	PCT-US94-10151A-64	Sequence 64, Appl
41	562	17.5	477	1	US-08-136-922-2	Sequence 2, Appl
42	471.5	14.6	634	3	US-09-041-232-2	Sequence 2, Appl
43	471.5	14.6	634	4	US-09-771-467C-2	Sequence 2, Appl
44	471.5	14.6	666	3	US-09-240-410-2	Sequence 2, Appl
45	406.5	12.6	606	3	US-09-041-236-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-556-422A-2
Sequence 2, Application US/08556422A

Patent No. 6576754

GENERAL INFORMATION:

APPLICANT: HALL, Kathryn T.

APPLICANT: FREEMAN, Gordon J.

APPLICANT: SCHULTZE, Joachim L.

APPLICANT: BOUSIOTIS, Vassiliaki

APPLICANT: NADLER, Lee M.

TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES

FILE REFERENCE: DRN-005CRA2

CURRENT APPLICATION NUMBER: US/08/556,422A

CURRENT FILING DATE: 1995-11-09

NUMBER OF SEQ ID NOS: 7

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 2

LENGTH: 862

TYPE: PRT

ORGANISM: Homo sapiens

US-08-556-422A-2

Query Match	Similarity	Score	Length	DB
Best Local	46.0%	Pred. No. 5.4e-126	862	
Matches	264	Conservative	89	Mismatches 188; Indels 33; Gaps 14;
28	VPRKTVSSGELATVVRFSQTCIGDPLFTLPTGLYVGARALFAF-SWEALELQGA	86		
27	IPRTWREHVLV--QHEPDIVYSALLSEDDITLIGAREAVFVNALNISKEQHE	84		
87	ISWEAPVEKTECIQKGNKNNQTECFNTRFLQPIYNASHLYVCGTAFQPKCTYVNLFTT	146		
85	YVMKVEDSKKAKCAKSKQTECLNTRVLOPISATSLYCGTNAFQPADCHLNLTSFK	144		
147	LEHGFEDGKGCPRDPKAGHAGLYDQGLSATLNANLGLPEPTLLRNKMGPHSKTEYL	206		
145	F-LGKNEGKGCPRDPKAGHAGLYDQGLSATLNANLGLPEPTLLRNKMGPHSKTEYL	202		
207	AFMLNEPFGVAGVAYVESVGSFTGDDKYVFFFRAYVSPDQYAGVAVARVYKGMG	266		
203	IPMLNEPFGVADYVTRKSPDSFGDDKYVFFFRAYVSPDQYAGVAVARVYKGMG	262		
267	GARTLQRTWTFELKARLACAPAMQLYFNOLAMHTLQDTSHANTFFGVQAGMDYL	326		
263	GLRTLOKKTWTFELKARLACAPAMQLYFNOLAMHTLQDTSHANTFFGVQAGMDYL	322		
327	SNICEYOLELQRTWTFELKARLACAPAMQLYFNOLAMHTLQDTSHANTFFGVQAGMDYL	382		
323	SAVCAVYNLSTAEVESHKRYMSTVTEQSHHKWVYNGVPRGACIDSEARAANYTS	382		

```

OY      383  SLEPDNLINLFVKKKPRPLBEOVGPMKSRLLYKKGNFPHILVADRYTGJLGDATYVTLFG 44.2
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      383  SLNLPDKLQLOVYKDHPLMDSDSTPTLDNRPLKDKDYNTQIIVDRQALDGSYTDVMEVS 44.2

         44.3  TGDGMLLAVALSGPVVHLIEELQLE--DQPMKSLVLSQSK--KLTFAGSRSQLVQLPVAD 49.9
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      44.3  TDRGALHRAISLEHAVHIIIEETQLQDPDFPVOTLLSSKKGRFVYAGNSGVQAPLAF 50.2

         OY      500  CIKRSCADCVLARDPYCAMSIVNTRCAV-----GGHFGSLLIQHMVTSIDTSGICN 55.1
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      503  CGKHGTCDCVLARDPYCAMSPEPTATVALHOTESPBSRG-----LIOE-MSGDAS-VCP 55.4

         OY      552  LRGSKIOSGPKKNITVVAQDVLVPCHLSSULA 58.5
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      555  DK-----SKGSTRQHFRRKHGCT-AELKCSQSKSULA 58.3

```

RESULT 2
US-08-556-422A-4
; Sequence 4, Application US/08556422A
Data: 085564

```

1      PATENT NO. 03/0734
2      GENERAL INFORMATION:
3      APPLICANT: HALL, Kathryn T.
4      APPLICANT: FREEMAN, Gordon J.
5      APPLICANT: SCHULTZE, Joachim L.
6      APPLICANT: BOUSSETIOTIS, Vassiliki
7      APPLICANT: NADLER, Lee M.
8      TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
9      FILE REFERENCE: DFN-005CPA2
10     CURRENT APPLICATION NUMBER: US/08/556,442A
11     CURRENT FILING DATE: 1995-11-09
12     NUMBER OF SEQ ID NOS: 7
13     SOFTWARE: FastSeq for Windows Version 4.0
14     SEQ ID NO 4:
15     LENGTH: 607
16     TYPE: PRT
17     ORGANISM: Mus musculus
18     US-08-556-442A-4

```

Query Match	35.7%	Score 1149;	DB 4;	length 607;
Best Local Similarity	43.6%	Pred. NO. 2.6e-118;		
Matches 247;	Conservative 89;	Mismatches 201;	Indels 30;	Gaps 11.

```

QY 41 VVRPSQTGIDFLTLTLTEPTGLLYVGAREALFAFSMALELOGA-----ISWEAVEK 96
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 LIRFEAEINISYATLALLSOGKTYLVGAREALFALNSLSLTPGGEYBELMSADNR 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 97 TEOLOKGNNOETECNFIREFLOPYNASHLYUCGTAFODKCTYYNMFLTLEHG----- 151
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 OQSGFKGKDPKRCNKYIKILLPLNSSHILTCGTAFAFSLCAYIHASFTLLADEGANYI 123
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 152 FEDGKRGKPYDPAKHAGLVDGELYSATLNFLTEPILTNMGPHHSMKTEYLAFWIN 211
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 124 LEDGKGKCPFDPNRSTALVYDGELETGTVSFOGDPAISSQSRRPT-KTESSLNMLO 1822
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 212 BPHFVGASVYBESVGSFTGDDDKYVFFERERAVESDCYAEQVAVARYARKCGDMGARTL 271
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 183 DPAFVASATSESLSPIGDDDKIYFFSETEGOEFEEFENTLYSRAVARYCKDGEGERYL 242
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 272 ORKTTFLKAKALGAPNMOLYFNOLQAMHTLODR--SHNHTTFEGYFOAQW--GDMYIS 327
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 243 QORMTSPFKAOLLCSPDGPFPFNVLQVFTLNPBQDKRLTISIGVFSOMHRTGTECS 302
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 328 AICEYQLEIQRVEGPKYKEHEEAOQMDRYDTPVPSPPGSCINNMHHRHGTSLELP 387
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 303 AICFTFMDVQKAPBGILTKKVRNREOQAMTTEHQVPTPFGACTTNSARERKTNSSLOLP 362
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 388 DNLINFVKRHLMEBOGVPWRMSRPLLVKKGTFNLVADRYTGLDQATYTVLETFGIDGM 447
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 363 DRYLNFLEKDHLEMDQOVR--SRLLTLDPBRYAQVAVARYPGLH-STYDYLEFLTGDDR 418
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 448 LTKVASLQPMWHLTFEOLDFDO--EPMRSLYVSQSKKLPAGSRSLQVQLPYADCIYRSC 506
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db 419 LKRAVTLSSRVHIIIEELQIFPGGPRVQNLLIDSHGLLIVASHSGVVQVPANCSLYPTC 478

Qy 507 ADCVLARDFPCAMSVTSCRCVANGGHPSGLIQHTSDTSG-----ICNLRGK----IQ 558
 ||||| ||| :
Db 479 GCDLARDFPCAMTSACRLASL---YQDLASRPWTODIGASVKELCKRSSITKANFLY 535

Qy 559 SGFAPKNITVYAGTDLVLPCHLSNLA 585
 ||| |||
Db 536 PGRFCCKOVQIQPNTVTNACLPLSNLA 562

RESULT 3
US-08-556-422A-3

```

Sequence 3, Application US/08556422A
Patent No. 6576754
GENERAL INFORMATION:
APPLICANT: HALL, Kathryn T.
APPLICANT: FREEMAN, Gordon J.
APPLICANT: SCHULTZE, Joachim L.
APPLICANT: BOUSSIOTIS, Vassiliki
APPLICANT: NADLER, Lee M.
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
FILE REFERENCE: DEN-005CPA2
CURRENT APPLICATION NUMBER: US/08/556,422A
CURRENT FILING DATE: 1995-11-09
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 655
TYPE: prf
ORGANISM: Homo sapiens
US-08-556-422A-3

```

Query Match 29.08; Score 932.5; DB 4; Length 655;

Best Local Similarity 38.4%; Pred. No. 3.5e-94;
Matches 208; Conservative 94; Mismatches 203; Indels 37; Gaps 15;

8 WL--LAARLMLGIGAEVMM-----NLVPRKTVSSGELATVVRFSQTGI--QDFLLTL 58

3 MTRIVCYLFWGVLITARANYQNGKNVPRKLSYKEMESNNVTTFENGLANSSSYHFFLL 62

59 TPTGTLVYGAREALFAFSMEALELQGAISWEAPVEKTECIQKGKNNQTECFNPIRELQ 118

b 63 DEERSHLYVGAKDHIFFSFDLVNINDQKIWVPVSYTRDECKWAGKDIKBCANFIKVLK 122

119 PYNASHLYVCCTYAFORCKTYVNM-----LFTLEHGEEDGKGKCPYDPAKGAGLLV 172

123 AYNQTHLYACGTGAFFHICTYIEIGHHPEDNIFKLENSHENGGRKSPYDPKLLTASLLI 182

173 DGEIYSATLNNEIQTETPIIRNMGPHSMKTE-YLAEFLNEPHFVGSAYVPESVGSETGD 231

183 DGELYSCTAEDFMGRDFAIFRTLGHHPIRTEQHDHSRWINDPKFISALHISESDNP---E 239

232 DDKVFEFERRAVESDCYAEQVAVRVARVCKGDMGGARTIQRKWTFPKARLACAP--- 288

240 DDKVYFFFRENAIDGESHGKATHARIGQICKNDFGSHSLVYNKWTTFKARLICSVGP 299

289 NWOLYFNOLAMHTLODTSWHNTTFFGVFOAOMGMYLSAICEYQLEEIQRVFEQPYKEY 348

300 GIDTDFEQLDQVFLMAFKDPKKNPVYVGVFTTSSNIEFGSAVCMYSMSDVRVFLGPIYAHK 359

349 HEEAQKWDRTDVPSPRPGSCINWHRHGYTSLIPDNLNFKKHPLMEEQVPRW 408

360 DGPNTQWVYQGRVPIPRPGTCSKTF--GGEIDSTKDLPPDVITFASSHPRAMINPYEPMN 417

409 SRPPLVKKGTN--ETHLVADRVTCGDGATYFWLFIGTGDGWLKAVSL--GPMVH----L 460

b
418 NRPIVKTQVNIQFTQIVYDRAVDAEDG-QIDVMEIGTDVGYVLKVVSIPIKETWYDLEEVL 476

461 IEELELDOEP--MRSVLVSQSKLLEFAGSRQVLQLPVADCIKY-RSCADCVILADPYC 517

477 LEEMTVF-REPTAISANELSTKQQLYIGSTAGVAQLPLHRCDIYGACAECCLARDPYC 535

OY 518 AM 519
Db 536 AM 537

RESULT 4

US-08-121-713D-54
; Sequence 54, Application US/08121713D
; Patent No. 5639856
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matches, David R.
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713D
; FILING DATE: 13-SEP-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELETYPE:
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-121-713D-54

Query Match 29.0%; Score 932.5; DB 1; Length 771;
Best Local Similarity 38.4%; Pred. No. 4.6e-94;
Matches 208; Conservative 94; Mismatches 203; Indels 37; Gaps 15;

OY 8 WL--LAARLWGLGIGAEYWM-----NLVPRKTVSSGELATVVRRSQTGI---QDFLTTL 58
Db 3 WLFRIVCLFNGVLLLTARANYONGKRNWRLSTYKEMESNNVITFNFNLAASSSTHTFL 62
OY 59 TEPFGLVYGAREALFAPSMELALOGAISWEAVEKTEIOGKNNQTECFNPIRFLQ 118
Db 63 DEESRLTYGAKDIIFSFDLVINIKDFOKITVPVSTTRDECKWAGKDKLKECANFIKVLK 122
OY 119 PYNASHLYVCGTVAFOPKCTIVNM-----LFTLEHGEFEDGKCKPYDPAKGAGLAV 172
Db 123 AYNQTHLYAGCTGAFHPICTYIEIGHPEDNIFKLENSHFNGSKSPYDPRKLTASLLI 182
OY 173 DGEIYSATLNFLETPEIILANMGRPHSMKTE-YLATVLANPHVGSAYVESVGSFTGD 231
Db 183 DGEIYSGTAADFEMGRDAIFRTLGHHPIRTECHDSRLNDPKFISAHLLIESDNP---E 239
OY 232 DDKYFFPEREAVSDCYAEQVAVARVAVYCGDKMGARTLQKMTTFPLKARLACAP--- 288
Db 240 DDKYFFPEREAVSDCYAEQVAVARVAVYCGDKMGARTLQKMTTFPLKARLACAP--- 299

OY 289 NWLYENOLAMHTIQTSTWNTFFGVQAOMQMDYLSAICEYQLEEIQRVEGPKREY 348
Db 300 GIDTFDELQDVFLNFKDPKPNVYGVFTTSSNIFKGSAYCMVSMDSRVRVFLGPVYHR 359
OY 349 HEEAKMDRYTDPVSPRRPGSCINMHRRCYTSLSLEPDLILNFVYKHPHMEQVGRW 408
Db 360 DGPYQWVYQGRVYPRPGTSPSKTF--GGFDSSTKDLDPDVIYTFARSHPMYRPVPMN 417
OY 409 SRPLLVRKGTN--FTHLVADRVTGIGATYTVLTFGTGDMILKAVSL--GPNVH---L 460
Db 418 NRPVITKTDVYKQFTQIYVDVDAEDG-QYDMFQIGTGVYLVAVSIFPRTWTDLEVYL 476
OY 461 IEEQLFQDEP--MRSVLTSQSKLLFNGSRSLQVLPVADCIKY-RSCADCVIARDPYC 517
Db 477 LEEWTVF-REPTAISAMELSTKQOOLYIGSTAGVAQLPLHRCDIYGRKACACCLARDPYC 535
OY 518 AM 519
Db 536 AM 537

RESULT 5

US-08-835-268-54
; Sequence 54, Application US/08835268
; Patent No. 5807826
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matches, David R.
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELETYPE:
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-835-268-54

Query Match 29.0%; Score 932.5; DB 1; Length 771;
Best Local Similarity 38.4%; Pred. No. 4.6e-94;
Matches 208; Conservative 94; Mismatches 203; Indels 37; Gaps 15;

OY 8 WL--LAARLWGLGIGAEYWM-----NLVPRKTVSSGELATVVRRSQTGI---QDFLTTL 58
Db 3 WLFRIVCLFNGVLLLTARANYONGKRNWRLSTYKEMESNNVITFNFNLAASSSTHTFL 62
OY 59 TEPFGLVYGAREALFAPSMELALOGAISWEAVEKTEIOGKNNQTECFNPIRFLQ 118
Db 63 DEESRLTYGAKDIIFSFDLVINIKDFOKITVPVSTTRDECKWAGKDKLKECANFIKVLK 122
OY 119 PYNASHLYVCGTVAFOPKCTIVNM-----LFTLEHGEFEDGKCKPYDPAKGAGLAV 172
Db 123 AYNQTHLYAGCTGAFHPICTYIEIGHPEDNIFKLENSHFNGSKSPYDPRKLTASLLI 182
OY 173 DGEIYSATLNFLETPEIILANMGRPHSMKTE-YLATVLANPHVGSAYVESVGSFTGD 231
Db 183 DGEIYSGTAADFEMGRDAIFRTLGHHPIRTECHDSRLNDPKFISAHLLIESDNP---E 239
OY 232 DDKYFFPEREAVSDCYAEQVAVARVAVYCGDKMGARTLQKMTTFPLKARLACAP--- 288
Db 240 DDKYFFPEREAVSDCYAEQVAVARVAVYCGDKMGARTLQKMTTFPLKARLACAP--- 299

```

Db      3 WLIRYCLFWGVLTLRANYONGKNVPRKLKSYKEMLSNNVITNGLANSSSYHTFLL 62
OY      59 TEPTGLLYGAREALFAFSMEALELOGAISWEAPVEKTECILOKGNNOGTECFNFRIFO 118
Db      63 DEBSRLTYGADHIFSEFLVNIKDFOKIWPVSYTRDECKRAGDILKECANFTKVAK 122
OY      119 PYNASHLYVCGTYAFOPKCTYVNM-----LFTLEHGEFEDGKGCYPDPAKAGHALLV 172
Db      123 AVNOTHLVACGCAFPICITYIEIGHNRPDNIKLENSHFENGRGKSPYDPKLLTASLLI 182
OY      173 DEELYSATLNFLGTPEPILRNKGRPHSKTE-YLAFWLNERPFGVSATVPESVGSFTGD 231
Db      183 DEELISGTADFGNDFAIFRTLGHHPIRTDQHSRMLNDPFIASHLISESDNP---E 239
OY      232 DDKVYFFERARVESDCYAEQVAVARYAKGDMGARTLORWTFFLAKRLACAP--- 288
Db      240 DDKVYFFERENADGSHSKATHARIGQICKNDFGHRLVYKWTFFLAKRLACSVPGN 299
OY      289 NMQLYFNOLQAMHTLDDTSMHNTTFEGVPOAGMDYLSAICEYOLEEIOREYFEGPYKEY 348
Db      300 GIDTHDELQDVFLNFKDPKPNVYGVFTSSNIFKGSVAVCMYSMSDVRVFLGYPYAR 359
OY      349 HEEAOKWDRYTDVPSPRGSCINMHRHGYTSSLELDPDNLINFKKHPLEMOGYPRM 408
Db      360 DDPNTQWVPYQGRVPRPGRCTPSKTF--GGFSTIDLPDDVITTFARSHPRAMNPFPMN 417
OY      409 SRPLVKKGTN--FTHLVADRYTGLDGAITYVLFICTGDMWLKAVSL--GPMVH---L 460
Db      418 NRPYVKTQVNGFTQIYVDRVDAEDG-QDYVAFICTDVGTVLKVYSIKETMYDLEVL 476
OY      461 IEELQLEFDEP--MRSVLVSQSKLLFAGSRQOLVLPVADCIKY-RSCADCVLARDPYC 517
Db      477 LEMTVF-REPAISAMELSTKQOOLYIGSTAGVQLPLRCDIYKACAECCCLARDPYC 535
OY      518 AW 519
Db      536 AW 537

```

RESULT 6
 US-09-060-692-54
 ; Sequence 54, Application US/09060692
 ; Patent No. 5935865
 ; GENERAL INFORMATION:
 ; APPLICANT: Goodman, Corey S.
 ; APPLICANT: Kolodkin, Alex L.
 ; APPLICANT: Matthes, David
 ; APPLICANT: Bentley, David R.
 ; APPLICANT: O'Connor, Timothy
 ; TITLE OF INVENTION: The Semaphorin Gene Family
 ; NUMBER OF SEQUENCES: 100
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 ; STREET: 268 Bush Street, Suite 3200
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/060,692
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/121,713
 ; FILING DATE: 13-SEP-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Osman, Richard A.

```

; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
;
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-09-060-692-54
;
; Query Match
; Best local Similarity 38.4%; Pred. No. 4,6e-94;
; Matches 208; Conservative 94; Mismatches 203; Indels 37; Gaps 15;
;
; OY      8 WL--LAARLWGLGIGAEVNM---NLVPRKTVSSGGLATVVRRESOTGI---ODFLTLTL 58
; Db      3 WLIRYCLFWGVLTLRANYONGKNVPRKLKSYKEMLSNNVITNGLANSSSYHTFLL 62
; OY      59 TEPTGLLYGAREALFAFSMEALELOGAISWEAPVEKTECILOKGNNOGTECFNFRIFO 118
; Db      63 DEBSRLTYGADHIFSEFLVNIKDFOKIWPVSYTRDECKRAGDILKECANFTKVAK 122
; OY      119 PYNASHLYVCGTYAFOPKCTYVNM-----LFTLEHGEFEDGKGCYPDPAKAGHALLV 172
; Db      123 AVNOTHLVACGCAFPICITYIEIGHNRPDNIKLENSHFENGRGKSPYDPKLLTASLLI 182
; OY      173 DEELYSATLNFLGTPEPILRNKGRPHSKTE-YLAFWLNERPFGVSATVPESVGSFTGD 231
; Db      183 DEELISGTADFGNDFAIFRTLGHHPIRTDQHSRMLNDPFIASHLISESDNP---E 239
; OY      232 DDKVYFFERARVESDCYAEQVAVARYAKGDMGARTLORWTFFLAKRLACAP--- 288
; Db      240 DDKVYFFERENADGSHSKATHARIGQICKNDFGHRLVYKWTFFLAKRLACSVPGN 299
; OY      289 NMQLYFNOLQAMHTLDDTSMHNTTFEGVPOAGMDYLSAICEYOLEEIOREYFEGPYKEY 348
; Db      300 GIDTHDELQDVFLNFKDPKPNVYGVFTSSNIFKGSVAVCMYSMSDVRVFLGYPYAR 359
; OY      349 HEEAOKWDRYTDVPSPRGSCINMHRHGYTSSLELDPDNLINFKKHPLEMOGYPRM 408
; Db      360 DDPNTQWVPYQGRVPRPGRCTPSKTF--GGFSTIDLPDDVITTFARSHPRAMNPFPMN 417
; OY      409 SRPLVKKGTN--FTHLVADRYTGLDGAITYVLFICTGDMWLKAVSL--GPMVH---L 460
; Db      418 NRPYVKTQVNGFTQIYVDRVDAEDG-QDYVAFICTDVGTVLKVYSIKETMYDLEVL 476
; OY      461 IEELQLEFDEP--MRSVLVSQSKLLFAGSRQOLVLPVADCIKY-RSCADCVLARDPYC 517
; Db      477 LEMTVF-REPAISAMELSTKQOOLYIGSTAGVQLPLRCDIYKACAECCCLARDPYC 535
; OY      518 AW 519
; Db      536 AW 537

```

RESULT 7
 US-08-833-391-54
 ; Sequence 54, Application US/08833391
 ; Patent No. 6013781
 ; GENERAL INFORMATION:
 ; APPLICANT: Goodman, Corey S.
 ; APPLICANT: Kolodkin, Alex L.
 ; APPLICANT: Matthes, David
 ; APPLICANT: Bentley, David R.
 ; APPLICANT: O'Connor, Timothy
 ; TITLE OF INVENTION: The Semaphorin Gene Family
 ; NUMBER OF SEQUENCES: 100
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

350 EAAKNDRTDVPSPRGSCINNMHRHGTSSLELDPNLTNFKKHPLEMOYGRWS 409
180 QSHTKMYRYPGPRGACIDSEARAANTSSINLPDKTLOFVKHPHLLDDSTPTDN 239
410 RPLVKGTPNTHLVADRVGLDGAATYVLEIGTGDMGLKAVSIGPVAHLIEELOLE-D 468
240 RPKIKKDVAHTQIYVDRTQALDGTVDYDMFVSTDRGLMKAISLEHVAHLIEELOLE-D 299
469 QEPKRSVLVSQSK--KILFAGSRQVLQVADCIKRSACADCVLARDPYCAMSVTSRC 526
300 FEYQVOTLLSKKGNRRFYAGNSGVQAPLAFGKHOTCEDCVLARDPYCAMSPRTATC 359
527 VAV-----GGHFGSLIDHWTSDTSGICNRSKISGSPKPKNTTVAGTDLVPC 578
360 VALHQTESPRKG-----LIOE-MSGDAS-VCPRK-----SKGSYRHFPHKGT-AELKC 406
579 HLSNLA 585
407 SOKSNLA 413

RESULT 11

US-09-308-1798-1
; Sequence 1, Application US/093081798
; Patent No. 643669
; GENERAL INFORMATION:
; APPLICANT: INAGAKI, Shinobu
; APPLICANT: FURUYAMA, Tatsuo
; TITLE OF INVENTION: NOVEL SEMAPHORIN GENES (1)
; FILE REFERENCE: 0020-4562P
; CURRENT APPLICATION NUMBER: US/09/308, 1798
; CURRENT FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/JP97/04111
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: JAPAN 321068/1996
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-308-1798-1

Query Match 26.5%; Score 852.5; DB 4; Length 775;
Best Local Similarity 36.6%; Pred. No. 3.6e-85;
Matches 210; Conservative 86; Mismatches 216; Indels 61; Gaps 22;

1 MAPMAYMLAARLWGLGIGAEVW-----NLVPRKTVSSGELATVYRPF--SQTGIO 51
1 MAP--AGHIIITLLMGHLL--ELMTPGHSANPSYRRLRSLKELLELRKTSIFQSPGLFL 56
52 DFLTLTLEPTGLLYGAREALFAFSMEAL--ELQALISMEAPVKKTECTIOKRNQDEC 110
57 DLHMLLDEYQERLEFVGGRDLVYSLNLERVSDGYREIYMPSTAVKVECKMGK--DANEC 115
111 FNFRIFLOPYNASHLYCGTAFAPOKCTYVNM-----LFTLEHGEEDGKGCOPYPA 164
116 ANYIRLVHNYRTHLLCATGAFDPHCAFIVGHSSEPLFHLSSHSEBRGRCOPRPN 175
165 KGAAGLLVDGELYATLTNNLTGTEPILIRNKGPHHSKTEY-----LAFWLNEMPHVGSAY 220
176 SSFSTVLYGNELFAGLYSDYWGDSALFRSMGKLGHITTEHDERL--LKEPKFVGSYM 232
221 VPESVGSPTGDDVVYFFERERAVESDCYAEQVAVARVAVYCKGDMGAKRTLOKRTPTLK 280
233 IPDNEDR---DDNMKTFEETKALEAENNAHTTTRVGRLLCVNMGGORILLVNMSTYLK 289
281 ARLACAPAMW---OLYFNOLQAMHTLQDTSWHNTTFEGVFOAQMGDMYLSAICEYOLEEI 337
290 ARLVCSPVGMGIGITFYDELEDDVLLPTTRDPKKNVIRFLFTNTSINIFGHAVCYVHMSI 349
338 QRVTEGPT-----KEYHEAOKMDRTDVPSPRGSCINNMHRHG--YTSLELDPNI 390

350 REARNGPVAHKEGREYH-----WSLYEGKVPYPRGSCAS--KVGSKGTGTRDYPDA 401
391 LNFVKKHPLMBEOYGRPKRSRLVLRKG--TPTFLVADRVGLDGAATYVLEIGTGDMGL 448
402 IRFAMHPLMOPTRPKPVHKKPLVKTGKYLRQLADRVRAEDG--QDVLVIGTDTGIV 460
449 LKAVSL-----GPVYH--LIEELOLE-DOEPKRSVLVSQSKKILFAGSRQVLQVADCI 501
461 LKVTIYNQELNEMBEVILELOIFKDPAPILISMESISKROQLYIGSAMVAQVRPHHCD 520
502 KYRS-CADCVLARDPYCAM-SVNTSRCAVAGH 532
521 MYGSACADCCCLARDPYCAMDGISCSRYPTGAH 553

RESULT 12

US-09-254-594-6
; Sequence 6, Application US/09254594
; Patent No. 6566094
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru
; APPLICANT: KIKUCHI, Kaoru
; TITLE OF INVENTION: NOVEL SEMAPHORIN GENE: SEMAPHORIN Y
; FILE REFERENCE: 0020-4527P
; CURRENT APPLICATION NUMBER: US/09/254, 594
; CURRENT FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: tissue type: child brain
; NAME/KEY: misc.feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Identification Method: P for resulting peptide
US-09-254-594-6

Query Match 22.6%; Score 727.5; DB 4; Length 930;
Best Local Similarity 34.9%; Pred. No. 3.9e-71;
Matches 193; Conservative 91; Mismatches 192; Indels 77; Gaps 24;

51 QDLTLTLEPTGLLYGAREALFAFSMEAL-----ISWEAPVKKTECTIOKG 103
64 QRFLLNRT-----LTVAAARDHVSFDLQA--EEEGEGLVPKRYLWTR--QDVCKAVRG 115
104 KNNQTECFNFRIFLOPYNASHLYCGTAFAPOKCTYVNM--LFTLEHGEEDGKGCOPYDP 163
116 KLTD-ECYNTIRLVLPWDSOTLLACGINSFSPVCRSYG--ITSLOQESBELSGOACRPDA 173
164 AKGAAGLLVDGELYSATLTNNLTGTEPILIRNKGPHHSKTEY--EYLAFWLNEMPHVGSAYVP 222
174 TQSNVAIRAESLSVSATADPOASDAVYRSLGQPPRLRSKAYDSKMLREPHFOA----- 229
223 ESVGSFTGDDKVVYFFERERAVESDCYAEQVAVARVAVYCKGDMGCA--RTLOKRTPTFLKA 281
230 -----LEHGHVYFFEFREVSIEDARLKGVSRAVCKRDMGSPALDRHHTSFKL 283
282 RLACGAP--NMOLYFNOLQAMHTLQDTSWH--NTTFFGVFOAQMGDMYLSAICEYOLEEIOR 339
284 RLNCSPVSDSFTFYDVAL--TGPNVNLHGSALEFGVTTTGTNTSIPGSAVCAFLIDETER 341
340 VFEQPYKEHYHEAOKMDRT--DPVSPRGSCINNMHRHG-----YTSLELDPNLTN 392
342 GFEKGFRQSLDGAAMTVSDEDRVSPRGSCA-----GVGGAALFSSSRDLPTDVL 394
393 FVKKHPLMBEOYGRPKRSRLVLRKG--TPTFLVADRVGLDGAATYVLEIGTGDMGLK 450
395 FIKAHPLIDPAVPVPTHQPRLTL--TSRALLTOVAVDGMAPHSNITVNFSGSNDGTVLK 452

```

OY 451 AV-----SLGPNVHLIEELQFDQPMR-----SLVLSGSKLLFAGSRNQ 491
:      :      :      :      :      :      :      :      :      :
DB 453 VLPGRSGSGPPILLIEEDAY--SPARCSGKRTAQTARRIIGLEIDTEGHLEFVAFSGC 510
:      :      :      :      :      :      :      :      :      :
OY 492 LVQLPVADCIKRSAC-ADCVLARDPYCAMSVNTSR-CVAVGHHFSLLO-HVMNSDTSG 548
:      :      :      :      :      :      :      :      :      :
DB 511 IYVLPISRCARHACORCLASODPYCGW--HSSRCVDIIRSGGTVDVDAQNOESMERHG 568
:      :      :      :      :      :      :      :      :      :
OY 549 ICLNRSGKIQSGP 561
:      :      :      :      :      :      :      :      :      :
DB 569 DCGDGTGSGQSGP 581
:      :      :      :      :      :      :      :      :      :

RESULT 13
US-08-556-422A-6
: Sequence 6, Application US/08556422A
: Patent No. 6576754
: GENERAL INFORMATION:
: APPLICANT: HALL, Kathryn T.
: APPLICANT: FREEMAN, Gordon J.
: APPLICANT: SCHULZE, Joachim L.
: APPLICANT: BOUSSTIOTIS, Vassiliki
: APPLICANT: NADLER, Lee M.
: TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
: FILE REFERENCE: DFN-005CPA2
: CURRENT APPLICATION NUMBER: US/08/556,422A
: CURRENT FILING DATE: 1995-11-09
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: FASTSEQ for Windows version 4.0
: SEQ ID NO 6
: LENGTH: 295
: TYPE: PRT
: ORGANISM: Gallus gallus
US-08-556-422A-6

Query Match      22.4%; Score 719.5; DB 4; Length 295;
Best Local Similarity 48.5%; Pred. No. 4,2e-71;
Matches 143; Conservative 49; Mismatches 96; Indels 7; Gaps 4;

OY 232 DDKYFFERRERAVESDCYAEQYVAARVARYCKDGMGARTLQRRKTTFLKARLACSAAPMNO 291
:      :      :      :      :      :      :      :      :      :
DB 1 DDKIFFEFVEVSEYEFQCKLMIRIARCKRDGGLRLQKMTSFLKARLICITIPDN 60
:      :      :      :      :      :      :      :      :      :

OY 232 LYFNQAMHTLQDTSMHNTTFYVFOAGMDMTLSAICEYOLEEIQRYF-EGPYKE--- 347
:      :      :      :      :      :      :      :      :      :
DB 61 LFNFTINNYFTLSKSTLKEPVYGVFTPOLNNVGLSACAVNMLSAVEEFSGKYMGSAT 120
:      :      :      :      :      :      :      :      :      :

OY 348 YHEAOKDRYDTPVSPRPGSCINNMHRRHGYTSSLLELDNLLNPFVKKHPRLMEQVGR 407
:      :      :      :      :      :      :      :      :      :
DB 121 VEQSHTKWRTNGELPNRPPGACINNEARALNTVSLNLPDCTLQFVXDHPMLDSDVTPV 180
:      :      :      :      :      :      :      :      :      :

OY 408 WSRPLLVKKGTFNTHLADVATGDLGATYTVLFTGDSMLKAVSLGFWHLIEELQF 467
:      :      :      :      :      :      :      :      :      :
DB 101 GDRPLPLVRDVKYTIIVDVRAALNGTIYDVAFISTDQALHKAISTENGHIIIEETQLF 240
:      :      :      :      :      :      :      :      :      :

OY 468 DQ-EPHRSILVLS--OSKRLLFAGSRSLQVQLPVDACIKYRSCADCVLARDPYCAM 519
:      :      :      :      :      :      :      :      :      :
DB 241 PFEEPVQTLTLLSKSRRLYLAGSNSGVYQSPVAFCDYTTTCFDCVLARDPYCAM 295
:      :      :      :      :      :      :      :      :      :

RESULT 14
US-08-121-713D-58
: Sequence 58, Application US/08121713D
: Patent No. 5639856
: GENERAL INFORMATION:
: APPLICANT: Goodman, Corey S.
: APPLICANT: Kolodkin, Alex L.
: APPLICANT: Matthes, David
: APPLICANT: Bentley, David R.
: APPLICANT: O'Connor, Timothy
: TITLE OF INVENTION: The Semaphorin Gene Family
: NUMBER OF SEQUENCES: 100

```

```

: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
: STREET: 268 Bush Street, Suite 3200
: CITY: San Francisco
: STATE: CA
: COUNTRY: USA
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/121,713D
: FILING DATE: 13-SEP-1993
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Osaman, Richard A.
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: B94-002-1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415)343-4341
: TELEFAX: (415) 343-4342
: TELEX:
: INFORMATION FOR SEQ ID NO: 58:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 730 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-121-713D-58

```

```

Query Match      21.9%; Score 704.5; DB 1; Length 730;
Best Local Similarity 31.3%; Pred. No. 9,2e-69;
Matches 190; Conservative 107; Mismatches 245; Indels 65; Gaps 21;

OY 6 AVMILARLWAGIGIGIEVMMNLV-PRKYVSGGLATVRRFSOTGIGDTLTLTEPTGL 64
:      :      :      :      :      :      :      :      :      :
DB 4 ALVAVALLDM-VALHAAVAVNDVSPKMYOFG-ERVQHLGNSHDKFKLEKHNSL 61
:      :      :      :      :      :      :      :      :      :

OY 65 LVYGAREALFAVSMLEA--LOGAISEAPVEKTECIQKGNKNOTECFNFTRFLQPYNA 122
:      :      :      :      :      :      :      :      :      :
DB 62 L-VGARANIYVNSLRDLTEFTEDRIHMHSSGAHRELCYKGR-SEDCONTIRVLAKIDD 119
:      :      :      :      :      :      :      :      :      :

OY 123 SHLYVCGTYAFQPKCTYVNMLTFTLEHGEF-----EDGKCKPYDPAKAGLLVDGELY 177
:      :      :      :      :      :      :      :      :      :
DB 120 DRYLCIGTNAVYPLRCH-----YALKDGDYVEKEYEGRGLCPFDHDSNTAISGQLY 174
:      :      :      :      :      :      :      :      :      :

OY 178 SATLNFLGTEPTILRNMPHSMKTEYLAFMLNEHFPVGSAYVPESVSGFTGDDKYVF 237
:      :      :      :      :      :      :      :      :      :
DB 175 SATVADFSGTDLPLIR--GPLRTERSDLQ--LMAFNFNVTMEY-----NDTIF 220
:      :      :      :      :      :      :      :      :      :

OY 238 FFRERAVESDCYAEQYVAARVARYCKDGMGARTLQRRKTTFLKARLACSAAP-NMOLYXNO 296
:      :      :      :      :      :      :      :      :      :
DB 221 FFRERAVETINCKALISVAVARVCKHDKGPHOGDKMTSFLKSLNCSVPGDYPTFYNE 280
:      :      :      :      :      :      :      :      :      :

OY 297 LQAMHTLQDTSMHNTT---FFGVFOAGMDMTLSAICEYOLEEIQRYFEGPYKEHHEAQ 353
:      :      :      :      :      :      :      :      :      :
DB 281 IOSTDIIEGNAGGVYEKTIYVFTTPVNSIGSNAVCAFSMSILESFGPPKEDEOTMS 340
:      :      :      :      :      :      :      :      :      :

OY 354 KMDRYTD-PVSPRPGSCINNMHRRHGYTSSLLELDNLLNPFVKKHPRLMEQVGPMSRPL 412
:      :      :      :      :      :      :      :      :      :
DB 341 NMLAVPSLKVPEPRPGQVND-----SRTLDPVSNVFAKSHTLMDQAVPAFETRI 391
:      :      :      :      :      :      :      :      :      :

OY 413 LVKKGST--NFTHLVAD-RVTGLDGAITYVLFTGTDGMLLKAIVSLGAPVH-----LIE 462
:      :      :      :      :      :      :      :      :      :
DB 392 LIRISLQYRETKIADVQVQRTPDGAIVDLFTGTDGKIKALNSASFSSDPTVDSVYE 451
:      :      :      :      :      :      :      :      :      :

OY 463 ELQTFDQ-EPHRSILVLS-----SOSKRLLFAGSRSLQVQLPVDACIKYRSCADCVLARDPY 516
:      :      :      :      :      :      :      :      :      :
DB 452 ELQVLPVGVPMNLVVRNDGDSKLIVYSDEILAIKIRHSGSKITKRCRCVSLQDPY 511
:      :      :      :      :      :      :      :      :      :

OY 517 CAMSVNTRCAVAVGHFQSL---LIQVMTSDTSGICMLRSGKIQSGPKNITVVAQT 572
:      :      :      :      :      :      :      :      :      :

```

[illegible]

RESULT 15
US-08-835-268-58

; Sequence 58, Application US/08835268
; Patent No. 5907936

GENERAL INFORMATION:

Query Match	21.9%	Score 704.5	DB 1	Length 730
Best Local Similarity	31.3%	Pred. No. 9.2e-69		
Matches 190	Conservative 107	Mismatches 245	Indels 65	Gaps 21

```

Db      175 SATVADFSGTDLPIYR--GPLRTERSDLKQ--LNAFNFNVTMEY-----NDIFF 220
QY      238 PPREAAVESDCIAEOVAVANAVAVCCADMGCGARTLQKRTTFPLKARLACASAP-NMOLYFPND 236
Db      221 FFEREAAVEYINGCKATYISRKAVCRKDKPPIQGGGDRMTSFLKSRINCVPQGYPRYFNE 280
QY      297 LQAMTLODTSMHNTT---FFGVFOQMGDMYLSACEIQLEIEIOVFEFEPKYEHENA 353
Db      261 IOSTSDIIIEGATGGQVEKLYGVTFPTTAVNISIGCSAVCAFSKMSILESFPQPREOETMNS 340
QY      354 KMDRYTD-PVPSPRBGSCINNMHRRRQYTSLSLEPDNINLNFYKKRHPIMEEOVGPRMSRPI 412
Db      341 NMLAVPSLKVPRPRRQGVND-----SRTLPDVSVNFKRSHTLMDAENAPAFTRPI 391
QY      413 LYKKGTT-NFTHLVAD-RVTGLDGAITYTLFICTGDKMLKRAVSLGAPWYH-----LIE 462
Db      392 LIRISLOYRFTTIAVDQOVRTPDGKAYDVLFTGTDGKVIKALNSAFPOSSDPTVDSVIE 451
QY      463 ETQLFEDQ-EPKRSIYL-----SQSKILFPAFSRSQLQVLPVADCIYRSCADCVLARDPY 516
Db      452 ELQVLRPGCVFVANLVYVRMDGDDSKLYVVSDEIILAKLHRGSDKITMCRREVSIQDPY 511
QY      517 CAMSVNTRSCVAVGCHFGSL---LIQHYMTSDTSGICNLKRSKIOSGAPKXNITVAGT 572
Db      512 CAMDVNVELKCTAVGSPDMGAGRRRRIQNTISLEHKAAGCRQPTETIASVPPTOPTKSSG 571
QY      573 DLYLPCH 579
Db      572 DEYHSIH 578

```

Search completed: August 7, 2003, 13:51:21
Job time : 31 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2003, 13:50:00 ; Search time 54 Seconds

(without alignments)
1310.756 Million cell updates/sec

Title: US-10-002-050-14

Perfect score: 3219
Sequence: 1 MAPHAWMLAARLWGIG.....PCHLSNLALPDSPRESSV 596

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/FCI_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/FCIUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3217	99.9	596	14	US-10-003-152-14 Sequence 14, Appl
2	3217	99.9	596	15	US-10-002-050-14 Sequence 14, Appl
3	3217	99.9	596	15	US-10-002-304-14 Sequence 14, Appl
4	3115.5	96.8	590	14	US-10-003-152-12 Sequence 12, Appl
5	3115.5	96.8	590	15	US-10-002-050-12 Sequence 12, Appl
6	3115.5	96.8	590	15	US-10-002-304-12 Sequence 12, Appl
7	3113.5	96.7	833	15	US-10-149-819-4 Sequence 22, Appl
8	3107.5	96.5	624	14	US-10-003-152-22 Sequence 22, Appl
9	3107.5	96.5	624	15	US-10-002-050-22 Sequence 22, Appl
10	3107.5	96.5	624	15	US-10-002-304-22 Sequence 22, Appl
11	2720	84.5	834	15	US-09-759-1308-335 Sequence 35, Appl
12	2720	84.5	834	15	US-10-189-123-65 Sequence 65, Appl
13	2635	81.9	785	10	US-09-989-920-218 Sequence 218, Appl
14	1240	38.5	837	11	US-09-946-374-253 Sequence 253, Appl
15	1240	38.5	837	14	US-10-052-586-454 Sequence 454, Appl

16	1240	38.5	837	15	US-10-174-590-454	Sequence 454, App
17	1240	38.5	837	15	US-10-176-758-454	Sequence 454, App
18	1240	38.5	837	15	US-10-175-737-454	Sequence 454, App
19	1240	38.5	837	15	US-10-173-706-454	Sequence 454, App
20	1240	38.5	837	15	US-10-175-738-454	Sequence 454, App
21	1240	38.5	837	15	US-10-175-752-454	Sequence 454, App
22	1240	38.5	837	15	US-10-176-482-454	Sequence 454, App
23	1240	38.5	837	15	US-10-176-757-454	Sequence 454, App
24	1240	38.5	837	15	US-10-176-913-454	Sequence 454, App
25	1240	38.5	837	15	US-10-180-552-454	Sequence 454, App
26	1240	38.5	837	15	US-10-180-557-454	Sequence 454, App
27	1240	38.5	837	15	US-10-173-700-454	Sequence 454, App
28	1240	38.5	837	15	US-10-174-572-454	Sequence 454, App
29	1240	38.5	837	15	US-10-174-579-454	Sequence 454, App
30	1240	38.5	837	15	US-10-174-582-454	Sequence 454, App
31	1240	38.5	837	15	US-10-174-588-454	Sequence 454, App
32	1240	38.5	837	15	US-10-175-739-454	Sequence 454, App
33	1240	38.5	837	15	US-10-175-740-454	Sequence 454, App
34	1240	38.5	837	15	US-10-175-743-454	Sequence 454, App
35	1240	38.5	837	15	US-10-176-488-454	Sequence 454, App
36	1240	38.5	837	15	US-10-176-492-454	Sequence 454, App
37	1240	38.5	837	15	US-10-176-747-454	Sequence 454, App
38	1240	38.5	837	15	US-10-176-750-454	Sequence 454, App
39	1240	38.5	837	15	US-10-176-985-454	Sequence 454, App
40	1240	38.5	837	15	US-10-176-987-454	Sequence 454, App
41	1240	38.5	837	15	US-10-176-992-454	Sequence 454, App
42	1240	38.5	837	15	US-10-176-993-454	Sequence 454, App
43	1240	38.5	837	15	US-10-184-658-454	Sequence 454, App
44	1240	38.5	837	15	US-10-176-991-454	Sequence 454, App
45	1240	38.5	837	15	US-10-173-695-454	Sequence 454, App

ALIGNMENTS

RESULT 1
US-10-003-152-14
; Sequence 14, Application US/10003152
; Publication No. US20020151494A1
; GENERAL INFORMATION:
; APPLICANT: Shinketsu, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Yang, Meljia
; APPLICANT: Boldog, Ferenc
; APPLICANT: Herrmann, John
; TITLE OF INVENTION: No. US20020151494A1 Amino Acid Sequences for Human Semaphorin
; FILE REFERENCE: 15966-554 Cura-54 CON-512
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (562)
; OTHER INFORMATION: Xaa may be any amino acid
US-10-003-152-14

Query Match 99.9%; Score 3217; DB 14; Length 596;
Best Local Similarity 100.0%; Pred. No. 4e-118;
Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAPHAWMLAARLWGIGAEVMMNVLPKRVYSGELATVVRRSQGIQDFLTLTLE 60
DB 1 MAPHAWMLAARLWGIGAEVMMNVLPKRVYSGELATVVRRSQGIQDFLTLTLE 60

```
QY 61 PGGLLVGARREALFAFSMELALOGAISWEAPVEKTECIQKNNQTECFNFIRLQY 120
DB 61 PGGLLVGARREALFAFSMELALOGAISWEAPVEKTECIQKNNQTECFNFIRLQY 120
QY 121 NASHLYVCGTYAFOPRCCTVNNMLFTLEHGEFEDGKCPYDPAKHAGLLVDGELYSAT 180
DB 121 NASHLYVCGTYAFOPRCCTVNNMLFTLEHGEFEDGKCPYDPAKHAGLLVDGELYSAT 180
QY 181 LNNEFLCTEPIILRNMGPHHSMTETYLAFMLNEPHFVGSAYVPSVGSFTGDDDKYFFPR 240
DB 181 LNNEFLCTEPIILRNMGPHHSMTETYLAFMLNEPHFVGSAYVPSVGSFTGDDDKYFFPR 240
QY 241 ERAVESDCYAEQVAVANAVARCKGDMGAGARTLQKWTTFKLARLACAPMMQLYFNLOAM 300
DB 241 ERAVESDCYAEQVAVANAVARCKGDMGAGARTLQKWTTFKLARLACAPMMQLYFNLOAM 300
QY 301 HTLQDTSMTNTTFFGVFOAQMGDMYLSAICEYQLEIQRVFEQPYREYHEEAOKMDRYTD 360
DB 301 HTLQDTSMTNTTFFGVFOAQMGDMYLSAICEYQLEIQRVFEQPYREYHEEAOKMDRYTD 360
QY 361 PVPSPRPGSCINNMHRRHGYTSSLELPDNLNFKKHPMLMEQVGPWRMSRPLLKKGTFN 420
DB 361 PVPSPRPGSCINNMHRRHGYTSSLELPDNLNFKKHPMLMEQVGPWRMSRPLLKKGTFN 420
QY 421 THLVADRVGTGLDCAITTVLPFIGTGDGMLKAVSLGFWHLIEELQLFDOEPMSLYLSOS 480
DB 421 THLVADRVGTGLDCAITTVLPFIGTGDGMLKAVSLGFWHLIEELQLFDOEPMSLYLSOS 480
QY 481 KLLFAGSRQSOLVQLPVADCIKRYSCADCVLADPYCANSVNTSRCAVAGHGHSLLIOH 540
DB 481 KLLFAGSRQSOLVQLPVADCIKRYSCADCVLADPYCANSVNTSRCAVAGHGHSLLIOH 540
QY 541 VMTSDTSGICNLGSKIQSGPXPKNITVAGTDLVLPCHLSSNLALPDSNPESSV 596
DB 541 VMTSDTSGICNLGSKIQSGPXPKNITVAGTDLVLPCHLSSNLALPDSNPESSV 596

RESULT 2
US-10-002-050-14
; Sequence 14, Application US/10002050
; Publication No. US20030032095A1
; GENERAL INFORMATION:
; APPLICANT: Shinketsu, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corline
; APPLICANT: Yang, Meijia
; APPLICANT: Boldog, Ferenc
; APPLICANT: Herrmann, John
; TITLE OF INVENTION: No. US20030032095A1el Nucleic Acid Sequences Encoding Human Sema
; FILE REFERENCE: 15966-554 Cura-54 CON-S14
; CURRENT APPLICATION NUMBER: US/10/002,050
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (562)
; OTHER INFORMATION: Xaa may be any amino acid
US-10-002-050-14

Query Match 99.9%; Score 3217; DB 15; Length 596;
Best Local Similarity 100.0%; Pred. No. 4e-318;
Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MAPHAAVMLLAARLWGLGIGAEYWNMLVPRKTVSSSEELATVVRFSQGTQIDQFLTLTTE 60
DB 1 MAPHAAVMLLAARLWGLGIGAEYWNMLVPRKTVSSSEELATVVRFSQGTQIDQFLTLTTE 60
QY 61 PGGLLVGARREALFAFSMELALOGAISWEAPVEKTECIQKNNQTECFNFIRLQY 120
DB 61 PGGLLVGARREALFAFSMELALOGAISWEAPVEKTECIQKNNQTECFNFIRLQY 120
QY 121 NASHLYVCGTYAFOPRCCTVNNMLFTLEHGEFEDGKCPYDPAKHAGLLVDGELYSAT 180
DB 121 NASHLYVCGTYAFOPRCCTVNNMLFTLEHGEFEDGKCPYDPAKHAGLLVDGELYSAT 180
QY 181 LNNEFLCTEPIILRNMGPHHSMTETYLAFMLNEPHFVGSAYVPSVGSFTGDDDKYFFPR 240
DB 181 LNNEFLCTEPIILRNMGPHHSMTETYLAFMLNEPHFVGSAYVPSVGSFTGDDDKYFFPR 240
QY 241 ERAVESDCYAEQVAVANAVARCKGDMGAGARTLQKWTTFKLARLACAPMMQLYFNLOAM 300
DB 241 ERAVESDCYAEQVAVANAVARCKGDMGAGARTLQKWTTFKLARLACAPMMQLYFNLOAM 300
QY 301 HTLQDTSMTNTTFFGVFOAQMGDMYLSAICEYQLEIQRVFEQPYREYHEEAOKMDRYTD 360
DB 301 HTLQDTSMTNTTFFGVFOAQMGDMYLSAICEYQLEIQRVFEQPYREYHEEAOKMDRYTD 360
QY 361 PVPSPRPGSCINNMHRRHGYTSSLELPDNLNFKKHPMLMEQVGPWRMSRPLLKKGTFN 420
DB 361 PVPSPRPGSCINNMHRRHGYTSSLELPDNLNFKKHPMLMEQVGPWRMSRPLLKKGTFN 420
QY 421 THLVADRVGTGLDCAITTVLPFIGTGDGMLKAVSLGFWHLIEELQLFDOEPMSLYLSOS 480
DB 421 THLVADRVGTGLDCAITTVLPFIGTGDGMLKAVSLGFWHLIEELQLFDOEPMSLYLSOS 480
QY 481 KLLFAGSRQSOLVQLPVADCIKRYSCADCVLADPYCANSVNTSRCAVAGHGHSLLIOH 540
DB 481 KLLFAGSRQSOLVQLPVADCIKRYSCADCVLADPYCANSVNTSRCAVAGHGHSLLIOH 540
QY 541 VMTSDTSGICNLGSKIQSGPXPKNITVAGTDLVLPCHLSSNLALPDSNPESSV 596
DB 541 VMTSDTSGICNLGSKIQSGPXPKNITVAGTDLVLPCHLSSNLALPDSNPESSV 596

RESULT 3
US-10-002-304-14
; Sequence 14, Application US/10002304
; Publication No. US20030036185A1
; GENERAL INFORMATION:
; APPLICANT: Shinketsu, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corline
; APPLICANT: Yang, Meijia
; APPLICANT: Boldog, Ferenc
; APPLICANT: Herrmann, John
; TITLE OF INVENTION: Polynucleotides and polypeptides encoded thereby
; FILE REFERENCE: 15966-554 Cura-54 CON-S8
; CURRENT APPLICATION NUMBER: US/10/002,304
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (562)
; OTHER INFORMATION: Xaa may be any amino acid
US-10-002-304-14

Query Match 99.9%; Score 3217; DB 15; Length 596;
```

Best Local Similarity 100.0%; Pred. No. 4e-318;
Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 MAPHAAVLLAARLMLGIGAEVMMNLVPRKTVSSGELATVYRRFSOTGIDFLLTLTE 60
    |||||||
DB 1 MAPHAAVLLAARLMLGIGAEVMMNLVPRKTVSSGELATVYRRFSOTGIDFLLTLTE 60
OY 61 PTGLLYGAREALFAFSMEALELQGAISWEAPVEKTECIQKGNNOECFNFRLOPY 120
    |||||||
DB 61 PTGLLYGAREALFAFSMEALELQGAISWEAPVEKTECIQKGNNOECFNFRLOPY 120
OY 121 NASHLYCGTAAPOPKCTYVNMFLFTLEHGEFEDGKGCPRDPKAGAGLLYDELVSAT 180
    |||||||
DB 121 NASHLYCGTAAPOPKCTYVNMFLFTLEHGEFEDGKGCPRDPKAGAGLLYDELVSAT 180
OY 181 LNNEFGTEPIILRNMGPHHSKTEYLAFWLNEPHFVSAAVYVESGFTGDDDVYFFER 240
    |||||||
DB 181 LNNEFGTEPIILRNMGPHHSKTEYLAFWLNEPHFVSAAVYVESGFTGDDDVYFFER 240
OY 241 ERAVESDCYAEQVAVARAVAVCKGDMGARTLQRRMTFLKARLACSAPNMOLYFNOLQAM 300
    |||||||
DB 241 ERAVESDCYAEQVAVARAVAVCKGDMGARTLQRRMTFLKARLACSAPNMOLYFNOLQAM 300
OY 301 HTLQDTSNHNTEFFGVQAOQMDYLSAICGYOLEEIOREVEGPKYKEYHEBAQKMDRTD 360
    |||||||
DB 301 HTLQDTSNHNTEFFGVQAOQMDYLSAICGYOLEEIOREVEGPKYKEYHEBAQKMDRTD 360
OY 361 PVSPRPSCINNNHRRHGYTSLELDPNLIINFVKRHPIMEQVGPMSRPLVKKGTNF 420
    |||||||
DB 361 PVSPRPSCINNNHRRHGYTSLELDPNLIINFVKRHPIMEQVGPMSRPLVKKGTNF 420
OY 421 THLVADRYTGLDGAATYVLFITGIDGMLLKAVSLGPMWHLLEELQLEFQDEPMRSLVLSQS 480
    |||||||
DB 421 THLVADRYTGLDGAATYVLFITGIDGMLLKAVSLGPMWHLLEELQLEFQDEPMRSLVLSQS 480
OY 481 KKLIFAGRSQVLQPLVADCKIKYRSCADCVLARDPYCAWSVNTSCAVAGHGFSLIOH 540
    |||||||
DB 481 KKLIFAGRSQVLQPLVADCKIKYRSCADCVLARDPYCAWSVNTSCAVAGHGFSLIOH 540
OY 541 VMTSDTSGICNLRGSKIOSGPKNITVYVAGTDVLPCHLSSNLALPDSNPESSV 596
    |||||||
DB 541 VMTSDTSGICNLRGSKIOSGPKNITVYVAGTDVLPCHLSSNLALPDSNPESSV 596
```

RESULT 4

```
US-10-003-152-12
; Sequence 12, Application US/10003152
; Publication NO. US20020151494A1
; GENERAL INFORMATION:
; APPLICANT: Shinkels, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Yang, Meijia
; APPLICANT: Boldog, Ferenc
; APPLICANT: Herrmann, John
; TITLE OF INVENTION: No. US20020151494A1 Amino Acid Sequences for Human Semaphorin-7
; FILE REFERENCE: 15966-554 Cura-54 CON-S12
; CURRENT APPLICATION NUMBER: US/10/003,152
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-003-152-12
```

Query Match 96.8%; Score 3115.5; DB 14; Length 590;
Best Local Similarity 99.1%; Pred. No. 8.6e-308;

Matches 577; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

```
OY 1 MAPHAAVLLAARLMLGIGAEVMMNLVPRKTVSSGELATVYRRFSOTGIDFLLTLTE 60
    |||||||
DB 1 MAPHAAVLLAARLMLGIGAEVMMNLVPRKTVSSGELATVYRRFSOTGIDFLLTLTE 60
OY 61 PTGLLYGAREALFAFSMEALELQGAISWEAPVEKTECIQKGNNOECFNFRLOPY 120
    |||||||
DB 61 PTGLLYGAREALFAFSMEALELQGAISWEAPVEKTECIQKGNNOECFNFRLOPY 120
OY 121 NASHLYCGTAAPOPKCTYVNMFLFTLEHGEFEDGKGCPRDPKAGAGLLYDELVSAT 180
    |||||||
DB 121 NASHLYCGTAAPOPKCTYVNMFLFTLEHGEFEDGKGCPRDPKAGAGLLYDELVSAT 180
OY 181 LNNEFGTEPIILRNMGPHHSKTEYLAFWLNEPHFVSAAVYVESGFTGDDDVYFFER 240
    |||||||
DB 181 LNNEFGTEPIILRNMGPHHSKTEYLAFWLNEPHFVSAAVYVESGFTGDDDVYFFER 240
OY 241 ERAVESDCYAEQVAVARAVAVCKGDMGARTLQRRMTFLKARLACSAPNMOLYFNOLQAM 300
    |||||||
DB 241 ERAVESDCYAEQVAVARAVAVCKGDMGARTLQRRMTFLKARLACSAPNMOLYFNOLQAM 300
OY 301 HTLQDTSNHNTEFFGVQAOQMDYLSAICGYOLEEIOREVEGPKYKEYHEBAQKMDRTD 360
    |||||||
DB 301 HTLQDTSNHNTEFFGVQAOQMDYLSAICGYOLEEIOREVEGPKYKEYHEBAQKMDRTD 360
OY 361 PVSPRPSCINNNHRRHGYTSLELDPNLIINFVKRHPIMEQVGPMSRPLVKKGTNF 420
    |||||||
DB 361 PVSPRPSCINNNHRRHGYTSLELDPNLIINFVKRHPIMEQVGPMSRPLVKKGTNF 420
OY 421 THLVADRYTGLDGAATYVLFITGIDGMLLKAVSLGPMWHLLEELQLEFQDEPMRSLVLSQS 480
    |||||||
DB 421 THLVADRYTGLDGAATYVLFITGIDGMLLKAVSLGPMWHLLEELQLEFQDEPMRSLVLSQS 480
OY 481 KKLIFAGRSQVLQPLVADCKIKYRSCADCVLARDPYCAWSVNTSCAVAGHGFSLIOH 540
    |||||||
DB 481 KKLIFAGRSQVLQPLVADCKIKYRSCADCVLARDPYCAWSVNTSCAVAGHGFSLIOH 540
OY 541 VMTSDTSGICNLRGSKIOSGPKNITVYVAGTDVLPCHLSS 582
    |||||||
DB 541 VMTSDTSGICNLRGSKIOSGPKNITVYVAGTDVLPCHLSS 581
```

RESULT 5

```
US-10-002-050-12
; Sequence 12, Application US/10002050
; Publication NO. US20030032095A1
; GENERAL INFORMATION:
; APPLICANT: Shinkels, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Yang, Meijia
; APPLICANT: Boldog, Ferenc
; APPLICANT: Herrmann, John
; TITLE OF INVENTION: No. US20030032095A1 Nucleic Acid Sequences Encoding Human Se
; FILE REFERENCE: 15966-554 Cura-54 CON-S14
; CURRENT APPLICATION NUMBER: US/10/002,050
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-050-12
```

Query Match 96.8%; Score 3115.5; DB 15; Length 590;
Best Local Similarity 99.1%; Pred. No. 8.6e-308;
Matches 577; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

```

OY      1  MAPHAAVLLAARLWGIGIGAEVMMNVLPRKTVSSGELATVVRFSQTIODPFLTLE 60
Db      1  MAPHAAVLLAARLWGIGIGAEVMMNVLPRKTVSSGELATVVRFSQTIODPFLTLE 60
OY      61  PTGLLYGAREALFAFSMEALELQCAISWEAVEKTECTIOGKNNOTCFNFIREFLO 120
Db      61  PTGLLYGAREALFAFSMEALELQCAISWEAVEKTECTIOGKNNOTCFNFIREFLO 120
OY      121  NASHLYVCGTAFOPKCTYVNNLFTLLEGEFEDKGKCPYPAPAGHAGLVNDELVSAT 180
Db      121  NASHLYVCGTAFOPKCTYVNNLFTLLEGEFEDKGKCPYPAPAGHAGLVNDELVSAT 180
OY      181  LNNFLGTEPILIRNNGPHHSMKTEYLATWLNPHFVGSAYVBEVSGSTGGDDKYFFPR 240
Db      181  LNNFLGTEPILIRNNGPHHSMKTEYLATWLNPHFVGSAYVBEVSGSTGGDDKYFFPR 240
OY      241  ERAVESDCYAROVVARVAVRVCBGDGGARTLORKTTFLKAKLACAPAMWOLYFNLOAM 300
Db      241  ERAVESDCYAROVVARVAVRVCBGDGGARTLORKTTFLKAKLACAPAMWOLYFNLOAM 300
OY      301  HTLDDTSMHNTTFEGVFOAOWDMYLSAICEYOLEIQHFEGPYKEVHEEAOKMDRYTD 360
Db      301  HTLDDTSMHNTTFEGVFOAOWDMYLSAICEYOLEIQHFEGPYKEVHEEAOKMDRYTD 360
OY      361  PVSPRPSCINNHRHRHGYTSSLPLPNILNFVKHPLMEBOVGRMRGAPLLYKGTNF 420
Db      361  PVSPRPSCINNHRHRHGYTSSLPLPNILNFVKHPLMEBOVGRMRGAPLLYKGTNF 420
OY      421  THLVADRYTGLDGYTYVLFGTGDMTLKAVSLGSPWHLIELDLFDEPKRSVLVSGS 480
Db      421  THLVADRYTGLDGYTYVLFGTGDMTLKAVSLGSPWHLIELDLFDEPKRSVLVSGS 480
OY      481  KKLFPAGRSQVOLQPVVADCIKYRSCACQVLARDPYCASVMTSRCAVAGHGPGSLIIO 540
Db      481  KKLFPAGRSQVOLQPVVADCIKYRSCACQVLARDPYCASVMTSRCAVAGHGPGSLIIO 540
OY      541  VMTSDTSGICMLRGSKIQSGPYKKNITVYVAGTDLVLPCHLSS 582
Db      541  VMTSDTSGICMLRGSK - KVRPTPKNITVYVAGTDLVLPCHLSS 581

RESULT 6
US-10-002-304-12
? Sequence 12, Application US/10002304
? Publication No. US20030036185A1
? GENERAL INFORMATION:
? APPLICANT: Shinketsu, Richard
? APPLICANT: Fernandes, Elma
? APPLICANT: Vernet, Corine
? APPLICANT: Yang, Meijia
? APPLICANT: Boldog, Ferenc
? APPLICANT: Herrmann, John
? TITLE OF INVENTION: Polynucleotides and polypeptides encoded thereby
? FILE REFERENCE: 15966-554 Cura-54 CON-58
? CURRENT FILING DATE: 2001-11-02
? PRIOR APPLICATION NUMBER: 09/604,286
? PRIOR FILING DATE: 2000-06-22
? PRIOR APPLICATION NUMBER: 60/140,584
? PRIOR FILING DATE: 1999-06-23
? NUMBER OF SEQ ID NOS: 49
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 12
? LENGTH: 590
? TYPE: PRT
? ORGANISM: Homo sapiens
US-10-002-304-12

```

Query Match	96.8%;	Score 3115.5;	DB 15;	Length 590;
Best Local Similarity	99.1%;	Pred. No. 8.6e-308;		
Matches 577; Conservative	1;	Mismatches 3;	Indels 1;	Gaps 1

QY	1	MAPHAANVLARLWLGIGAEVMNVLVRKTVSSSELTAVYRFQOTGIDQDLTLTITE	60
Db	1	MAPHAANVLARLWLGIGAEVMNVLVRKTVSSSELTAVYRFQOTGIDQDLTLTITE	60
QY	61	PTGLLYGAREALFAFSMEALEOGAISWEAPVEKTECIQKKNQOTECFNFIRELOPY	120
Db	61	PTGLLYGAREALFAFSMEALEOGAISWEAPVEKTECIQKKNQOTECFNFIRELOPY	120
QY	121	NASHLYVCGTAFQORCTYVNNLPTFLLENGEEDGCKCPYDPKAKHAGLVDGELYSAT	180
Db	121	NASHLYVCGTAFQORCTYVNNLPTFLLENGEEDGCKCPYDPKAKHAGLVDGELYSAT	180
QY	181	LNNFLGTEPILIRNNGPHHSMTKEYLAFMLNPFHFYGSAYVPEVSQSFSGDDDKYFFFR	240
Db	181	LNNFLGTEPILIRNNGPHHSMTKEYLAFMLNPFHFYGSAYVPEVSQSFSGDDDKYFFFR	240
QY	241	ERAVESDCYAEQVAVARVARVCKDGMGARTLQKMTPTFLKARLACSA PMNQLYFNLOAM	300
Db	241	ERAVESDCYAEQVAVARVARVCKDGMGARTLQKMTPTFLKARLACSA PMNQLYFNLOAM	300
QY	301	HTLODTSWNTTFFCGYFQAGMDMTLSAICEYOLEIQRVFBSPRYEYHEEAOKMRYTD	360
Db	301	HTLODTSWNTTFFCGYFQAGMDMTLSAICEYOLEIQRVFBSPRYEYHEEAOKMRYTD	360
QY	361	PVSPRPGCINNMRHRGCTSSLEPDNITLNFVKHPLMEBOVGRMSPLLYRKGTNF	420
Db	361	PVSPRPGCINNMRHRGCTSSLEPDNITLNFVKHPLMEBOVGRMSPLLYRKGTNF	420
QY	421	THLVADRYTGLDGAITYVLFIGTGDGMLLKAVSLGPMVHLIEELOLFDQEPMSLVLSOS	480
Db	421	THLVADRYTGLDGAITYVLFIGTGDGMLLKAVSLGPMVHLIEELOLFDQEPMSLVLSOS	480
QY	481	KLLEFAGSLSQVLQPVADICIRYSGADCVLARDPCCAMSVTNSRCVYANGHGSLIOH	540
Db	481	KLLEFAGSLSQVLQPVADICIRYSGADCVLARDPCCAMSVTNSRCVYANGHGSLIOH	540
QY	541	WTSQTSGLICNLRGSKIOGSPKNTTYAAGDVLVPCHSS	582
Db	541	WTSQTSGLICNLRGSK - KVRPTPKNTTYAAGDVLVPCHSS	581

```

RESULT 7
US-10-149-819-4
: Sequence 4, Application US/10149819
: Publication No. US20030044913A1
: GENERAL INFORMATION:
: APPLICANT: INCYTE GENOMICS, INC.
: APPLICANT: YDE, Henry
: APPLICANT: AZIMZAI, Yalda
: APPLICANT: TANG, Y. Tom
: APPLICANT: PATTERSON, Chandra
: APPLICANT: BAUGHN, Mariah R.
: APPLICANT: LU, Dying Alma M.
: APPLICANT: SHAH, Puri
: APPLICANT: LAL, FreeLi
: APPLICANT: AU-YOUNG, Janice
: APPLICANT: BUREFORD, Neil
: TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
: FILE REFERENCE: PE-0760 PCT
: CURRENT APPLICATION NUMBER: US/10/149, 819
: CURRENT FILING DATE: 2002-06-10
: PRIOR APPLICATION NUMBER: 60/172,882; 60/172,354
: PRIOR FILING DATE: 1999-12-10; 1999-12-16
: NUMBER OF SEQ ID NOS: 42
: SOFTWARE: PERL Program
: SEQ ID NO 4
: LENGTH: 833
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: incyte ID No. US20030044913A1 1674661CD1
US-10-149-819-4

```

```

Query Match      96.7%; Score 3113.5; DB 15; Length 833;
Best Local Similarity 96.3%; Pred. No. 2,3e-307;
Matches 579; Conservative 4; Mismatches 7; Indels 11; Gaps 2;

QY 1 MAPHMAVWLLAARLWGIGIGAEVMMNLVPRKTVSSGELATVYRRFSQTGIDDFLTITTE 60
   1 MAPHMAVWLLAARLWGIGIGAEVMMNLVPRKTVSSGELATVYRRFSQTGIDDFLTITTE 60
DB 1 MAPHMAVWLLAARLWGIGIGAEVMMNLVPRKTVSSGELATVYRRFSQTGIDDFLTITTE 60
QY 61 PTGLLYGAREALFAFSHEALELQGAISWEAPVEKTECIQKGNKNOTECNFIFRLQPY 120
   61 PTGLLYGAREALFAFSHEALELQGAISWEAPVEKTECIQKGNKNOTECNFIFRLQPY 120
DB 61 PTGLLYGAREALFAFSHEALELQGAISWEAPVEKTECIQKGNKNOTECNFIFRLQPY 120
QY 121 NASHLYCGTYAFQPKCTYVNMFLTLEHGEFEDGKGCPRDPAKGHAGLLVDEGLYSAT 180
   121 NASHLYCGTYAFQPKCTYVNMFLTLEHGEFEDGKGCPRDPAKGHAGLLVDEGLYSAT 180
DB 121 NASHLYCGTYAFQPKCTYVNMFLTLEHGEFEDGKGCPRDPAKGHAGLLVDEGLYSAT 180
QY 181 LNNFLGTEPILIRNKGPHHSKTEYLAFWLNEPHFVGSAYVPESVGSFTGDDKYFFFR 240
   181 LNNFLGTEPILIRNKGPHHSKTEYLAFWLNEPHFVGSAYVPESVGSFTGDDKYFFFR 240
DB 181 LNNFLGTEPILIRNKGPHHSKTEYLAFWLNEPHFVGSAYVPESVGSFTGDDKYFFFR 240
QY 241 ERAVESDCYAEQVAVARAVKCGDMGARTLQKRTTFLKARLACSAFNMOLYFNQLOAM 300
   241 ERAVESDCYAEQVAVARAVKCGDMGARTLQKRTTFLKARLACSAFNMOLYFNQLOAM 300
DB 241 ERAVESDCYAEQVAVARAVKCGDMGARTLQKRTTFLKARLACSAFNMOLYFNQLOAM 300
QY 301 HTLQDTSMHNTTFEGVFOAQWDMYLSAICEYOLEIQRVEGPKYKEYHEEAOKMDRYTD 360
   301 HTLQDTSMHNTTFEGVFOAQWDMYLSAICEYOLEIQRVEGPKYKEYHEEAOKMDRYTD 360
DB 301 HTLQDTSMHNTTFEGVFOAQWDMYLSAICEYOLEIQRVEGPKYKEYHEEAOKMDRYTD 360
QY 361 PVPSPRGSCINNNHRRHGYTSSLELPRNIIINFVKKHPIMEOVGPRKSRPLVYKKGTF 420
   361 PVPSPRGSCINNNHRRHGYTSSLELPRNIIINFVKKHPIMEOVGPRKSRPLVYKKGTF 420
DB 361 PVPSPRGSCINNNHRRHGYTSSLELPRNIIINFVKKHPIMEOVGPRKSRPLVYKKGTF 420
QY 421 THLVADRVGTGDAATYTVLFTGTGDMLLKAVSLGPMVHLLEIQLDFQEPERSLYLSOS 480
   421 THLVADRVGTGDAATYTVLFTGTGDMLLKAVSLGPMVHLLEIQLDFQEPERSLYLSOS 480
DB 421 THLVADRVGTGDAATYTVLFTGTGDMLLKAVSLGPMVHLLEIQLDFQEPERSLYLSOS 480
QY 481 KLLFAGRSQSLVOLPVADCIRKYSACDVCYLARDPYCAMSVTSCVAVGSHGSLLIQH 540
   481 KLLFAGRSQSLVOLPVADCIRKYSACDVCYLARDPYCAMSVTSCVAVGSHGSLLIQH 540
DB 481 KLLFAGRSQSLVOLPVADCIRKYSACDVCYLARDPYCAMSVTSCVAVGSHGSLLIQH 540
QY 541 VMTSDTSGICNLKRGSK-KVRPTPKNTIYVAGTDVLVPCHLSSNLHAHWTEGGRDLPAEQ 599
   541 VMTSDTSGICNLKRGSK-KVRPTPKNTIYVAGTDVLVPCHLSSNLHAHWTEGGRDLPAEQ 599
DB 541 VMTSDTSGICNLKRGSK-KVRPTPKNTIYVAGTDVLVPCHLSSNLHAHWTEGGRDLPAEQ 599
QY 591 P 591
   591 P 591
DB 600 P 600
   600 P 600

RESULT 8
US-10-003-152-22
: Sequence 22, Application US/10003152
: Publication No. US20020151494A1
: GENERAL INFORMATION:
: APPLICANT: Shimkets, Richard
: APPLICANT: Fernandes, Elma
: APPLICANT: Vernet, Corine
: APPLICANT: Yang, Meljia
: APPLICANT: Boldog, Ferenc
: APPLICANT: Herrmann, John
: TITLE OF INVENTION: No. US20020151494A1 Amino Acid Sequences for Human Semaphorin-1
: FILE REFERENCE: 15966-554 Cura-54 CON-S12
: CURRENT APPLICATION NUMBER: US/10/003,152
: CURRENT FILING DATE: 2001-11-02
: PRIOR APPLICATION NUMBER: 09/604,286
: PRIOR FILING DATE: 2000-06-22
: PRIOR APPLICATION NUMBER: 60/140,584
: PRIOR FILING DATE: 1999-06-23
: NUMBER OF SEQ ID NOS: 49
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 22
: LENGTH: 624

```

```

: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-003-152-22
Query Match      96.5%; Score 3107.5; DB 14; Length 624;
Best Local Similarity 99.0%; Pred. No. 6.1e-307;
Matches 576; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 1 MAPHMAVWLLAARLWGIGIGAEVMMNLVPRKTVSSGELATVYRRFSQTGIDDFLTITTE 60
   1 MAPHMAVWLLAARLWGIGIGAEVMMNLVPRKTVSSGELATVYRRFSQTGIDDFLTITTE 60
DB 1 MAPHMAVWLLAARLWGIGIGAEVMMNLVPRKTVSSGELATVYRRFSQTGIDDFLTITTE 60
QY 61 PTGLLYGAREALFAFSHEALELQGAISWEAPVEKTECIQKGNKNOTECNFIFRLQPY 120
   61 PTGLLYGAREALFAFSHEALELQGAISWEAPVEKTECIQKGNKNOTECNFIFRLQPY 120
DB 61 PTGLLYGAREALFAFSHEALELQGAISWEAPVEKTECIQKGNKNOTECNFIFRLQPY 120
QY 121 NASHLYCGTYAFQPKCTYVNMFLTLEHGEFEDGKGCPRDPAKGHAGLLVDEGLYSAT 180
   121 NASHLYCGTYAFQPKCTYVNMFLTLEHGEFEDGKGCPRDPAKGHAGLLVDEGLYSAT 180
DB 121 NASHLYCGTYAFQPKCTYVNMFLTLEHGEFEDGKGCPRDPAKGHAGLLVDEGLYSAT 180
QY 181 LNNFLGTEPILIRNKGPHHSKTEYLAFWLNEPHFVGSAYVPESVGSFTGDDKYFFFR 240
   181 LNNFLGTEPILIRNKGPHHSKTEYLAFWLNEPHFVGSAYVPESVGSFTGDDKYFFFR 240
DB 181 LNNFLGTEPILIRNKGPHHSKTEYLAFWLNEPHFVGSAYVPESVGSFTGDDKYFFFR 240
QY 241 ERAVESDCYAEQVAVARAVKCGDMGARTLQKRTTFLKARLACSAFNMOLYFNQLOAM 300
   241 ERAVESDCYAEQVAVARAVKCGDMGARTLQKRTTFLKARLACSAFNMOLYFNQLOAM 300
DB 241 ERAVESDCYAEQVAVARAVKCGDMGARTLQKRTTFLKARLACSAFNMOLYFNQLOAM 300
QY 301 HTLQDTSMHNTTFEGVFOAQWDMYLSAICEYOLEIQRVEGPKYKEYHEEAOKMDRYTD 360
   301 HTLQDTSMHNTTFEGVFOAQWDMYLSAICEYOLEIQRVEGPKYKEYHEEAOKMDRYTD 360
DB 301 HTLQDTSMHNTTFEGVFOAQWDMYLSAICEYOLEIQRVEGPKYKEYHEEAOKMDRYTD 360
QY 361 PVPSPRGSCINNNHRRHGYTSSLELPRNIIINFVKKHPIMEOVGPRKSRPLVYKKGTF 420
   361 PVPSPRGSCINNNHRRHGYTSSLELPRNIIINFVKKHPIMEOVGPRKSRPLVYKKGTF 420
DB 361 PVPSPRGSCINNNHRRHGYTSSLELPRNIIINFVKKHPIMEOVGPRKSRPLVYKKGTF 420
QY 421 THLVADRVGTGDAATYTVLFTGTGDMLLKAVSLGPMVHLLEIQLDFQEPERSLYLSOS 480
   421 THLVADRVGTGDAATYTVLFTGTGDMLLKAVSLGPMVHLLEIQLDFQEPERSLYLSOS 480
DB 421 THLVADRVGTGDAATYTVLFTGTGDMLLKAVSLGPMVHLLEIQLDFQEPERSLYLSOS 480
QY 481 KLLFAGRSQSLVOLPVADCIRKYSACDVCYLARDPYCAMSVTSCVAVGSHGSLLIQH 540
   481 KLLFAGRSQSLVOLPVADCIRKYSACDVCYLARDPYCAMSVTSCVAVGSHGSLLIQH 540
DB 481 KLLFAGRSQSLVOLPVADCIRKYSACDVCYLARDPYCAMSVTSCVAVGSHGSLLIQH 540
QY 541 VMTSDTSGICNLKRGSK-KVRPTPKNTIYVAGTDVLVPCHLSS 581
   541 VMTSDTSGICNLKRGSK-KVRPTPKNTIYVAGTDVLVPCHLSS 581
DB 541 VMTSDTSGICNLKRGSK-KVRPTPKNTIYVAGTDVLVPCHLSS 581

RESULT 9
US-10-002-050-22
: Sequence 22, Application US/10002050
: Publication No. US20030032095A1
: GENERAL INFORMATION:
: APPLICANT: Shimkets, Richard
: APPLICANT: Fernandes, Elma
: APPLICANT: Vernet, Corine
: APPLICANT: Yang, Meljia
: APPLICANT: Boldog, Ferenc
: APPLICANT: Herrmann, John
: TITLE OF INVENTION: No. US20030032095A1 Nucleic Acid Sequences Encoding Human Se
: FILE REFERENCE: 15966-554 Cura-54 CON-S14
: CURRENT APPLICATION NUMBER: US/10/002,050
: CURRENT FILING DATE: 2001-11-02
: PRIOR APPLICATION NUMBER: 09/604,286
: PRIOR FILING DATE: 2000-06-22
: PRIOR APPLICATION NUMBER: 60/140,584
: PRIOR FILING DATE: 1999-06-23
: NUMBER OF SEQ ID NOS: 49
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 22
: LENGTH: 624
: TYPE: PRT

```

```

; ORGANISM: Homo sapiens
US-10-002-050-22
Query Match      96.5%; Score 3107.5; DB 15; Length 624;
Best Local Similarity 99.0%; Pred. No. 6.1e-307;
Matches 576; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 1 MAPHMAVWLLAARLWGLGIGAEYWMNLVPRKTVSSGELATVVRFSQGTGIDFLTTLTE 60
DB 1 MAPHMAVWLLAARLWGLGIGAEYWMNLVPRKTVSSGELATVVRFSQGTGIDFLTTLTE 60
OY 61 PTGLLVGAREALFAFMSMELELOGAISWEAPVEKTECIQKKNNOTECFNFIRLQAP 120
DB 61 PTGLLVGAREALFAFMSMELELOGAISWEAPVEKTECIQKKNNOTECFNFIRLQAP 120
OY 121 NASHLYVCGTYAFQPCCTVYNNMLFTLEHGEFEFGKGCYPDPKAGHAGLLVGGELYSAT 180
DB 121 NASHLYVCGTYAFQPCCTVYNNMLFTLEHGEFEFGKGCYPDPKAGHAGLLVGGELYSAT 180
OY 181 LNPFLETEPIILNNMGPBHSKTEYLAFWLNPEHFVGSAYVPSVGSFTGDDDKYFFFR 240
DB 181 LNPFLETEPIILNNMGPBHSKTEYLAFWLNPEHFVGSAYVPSVGSFTGDDDKYFFFR 240
OY 241 ERAVESDCYAEQVAVARVCKDGMGARTLQKRTTFLKARLACGAPNMQLYFNOLQAM 300
DB 241 ERAVESDCYAEQVAVARVCKDGMGARTLQKRTTFLKARLACGAPNMQLYFNOLQAM 300
OY 301 HTLQDTSMHNTTFEGVFOAQMGDMYLSAICEYOLEIQRVFEQPYKEYHEBAQKMDRYTD 360
DB 301 HTLQDTSMHNTTFEGVFOAQMGDMYLSAICEYOLEIQRVFEQPYKEYHEBAQKMDRYTD 360
OY 361 PVSPPRGSCINNMHRRHGTSSLELPDNLNFKVKKHPLMEBOVGRMSRPLLYKKGTFNR 420
DB 361 PVSPPRGSCINNMHRRHGTSSLELPDNLNFKVKKHPLMEBOVGRMSRPLLYKKGTFNR 420
OY 421 THLVADRYTGDSATYTVLEFIGTGDGMLKAVSLGPMVHLIEELQLEFDPKMSLYLSOS 480
DB 421 THLVADRYTGDSATYTVLEFIGTGDGMLKAVSLGPMVHLIEELQLEFDPKMSLYLSOS 480
OY 481 KLLFAGSRNSQLVQLPVADCIKTRSCADCVLARDPYCAMSVMTRCAVAGHSGSLILQH 540
DB 481 KLLFAGSRNSQLVQLPVADCIKTRSCADCVLARDPYCAMSVMTRCAVAGHSGSLILQH 540
OY 541 VMTSDTSGICNLNGSKIOSGPRKNTTVAGTDVLPCHLSS 582
DB 541 VMTSDTSGICNLNGSK-KVPRPKNTTVAGTDVLPCHLSS 581

RESULT 10
US-10-002-304-22
; Sequence 22, Application US/10002304
; Publication No. US20030036185A1
; GENERAL INFORMATION:
; APPLICANT: Shinkels, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Verneet, Corline
; APPLICANT: Yang, Meljia
; APPLICANT: Boldog, Ferenc
; APPLICANT: Hermann, John
; TITLE OF INVENTION: Polynucleotides and polypeptides encoded thereby
; FILE REFERENCE: 15966-554 Cura-54 CON-58
; CURRENT APPLICATION NUMBER: US/10/002,304
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```

US-10-002-304-22
Query Match      96.5%; Score 3107.5; DB 15; Length 624;
Best Local Similarity 99.0%; Pred. No. 6.1e-307;
Matches 576; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 1 MAPHMAVWLLAARLWGLGIGAEYWMNLVPRKTVSSGELATVVRFSQGTGIDFLTTLTE 60
DB 1 MAPHMAVWLLAARLWGLGIGAEYWMNLVPRKTVSSGELATVVRFSQGTGIDFLTTLTE 60
OY 61 PTGLLVGAREALFAFMSMELELOGAISWEAPVEKTECIQKKNNOTECFNFIRLQAP 120
DB 61 PTGLLVGAREALFAFMSMELELOGAISWEAPVEKTECIQKKNNOTECFNFIRLQAP 120
OY 121 NASHLYVCGTYAFQPCCTVYNNMLFTLEHGEFEFGKGCYPDPKAGHAGLLVGGELYSAT 180
DB 121 NASHLYVCGTYAFQPCCTVYNNMLFTLEHGEFEFGKGCYPDPKAGHAGLLVGGELYSAT 180
OY 181 LNPFLETEPIILNNMGPBHSKTEYLAFWLNPEHFVGSAYVPSVGSFTGDDDKYFFFR 240
DB 181 LNPFLETEPIILNNMGPBHSKTEYLAFWLNPEHFVGSAYVPSVGSFTGDDDKYFFFR 240
OY 241 ERAVESDCYAEQVAVARVCKDGMGARTLQKRTTFLKARLACGAPNMQLYFNOLQAM 300
DB 241 ERAVESDCYAEQVAVARVCKDGMGARTLQKRTTFLKARLACGAPNMQLYFNOLQAM 300
OY 301 HTLQDTSMHNTTFEGVFOAQMGDMYLSAICEYOLEIQRVFEQPYKEYHEBAQKMDRYTD 360
DB 301 HTLQDTSMHNTTFEGVFOAQMGDMYLSAICEYOLEIQRVFEQPYKEYHEBAQKMDRYTD 360
OY 361 PVSPPRGSCINNMHRRHGTSSLELPDNLNFKVKKHPLMEBOVGRMSRPLLYKKGTFNR 420
DB 361 PVSPPRGSCINNMHRRHGTSSLELPDNLNFKVKKHPLMEBOVGRMSRPLLYKKGTFNR 420
OY 421 THLVADRYTGDSATYTVLEFIGTGDGMLKAVSLGPMVHLIEELQLEFDPKMSLYLSOS 480
DB 421 THLVADRYTGDSATYTVLEFIGTGDGMLKAVSLGPMVHLIEELQLEFDPKMSLYLSOS 480
OY 481 KLLFAGSRNSQLVQLPVADCIKTRSCADCVLARDPYCAMSVMTRCAVAGHSGSLILQH 540
DB 481 KLLFAGSRNSQLVQLPVADCIKTRSCADCVLARDPYCAMSVMTRCAVAGHSGSLILQH 540
OY 541 VMTSDTSGICNLNGSKIOSGPRKNTTVAGTDVLPCHLSS 582
DB 541 VMTSDTSGICNLNGSK-KVPRPKNTTVAGTDVLPCHLSS 581

RESULT 11
US-09-759-130B-335
; Sequence 335, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackey, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodgar, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE REFERENCE: NP100-3550N1M
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
```



```

; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FASTSEQ for Windows version 4.0
; SEQ ID NO: 335
; LENGTH: 834
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-759-1308-335

```

```

Query Match      84.5%; Score 2720; DB 11; Length 834;
Best Local Similarity 83.1%; Pred. No. 2,5e-267;
Matches 501; Conservative 41; Mismatches 47; Indels 14; Gaps 4;

```

```

QY 1 MAPHAWMLAARLMGIGAGAEVMMNLYPRKTVSSGELATVYVRRFSQTGIDFILTILTE 60
DB 1 MAPHAWMLAARLMGIGAGAEVMMNLYPRKTVSSGELATVYVRRFSQTGIDFILTILTE 60
QY 61 PTGLLYGAREALPAFSEALELOGAISWEAPVEKTECIQKKNOTECFNFTRFLQPY 120
DB 61 HSGLLYGAREALPAFSEALELOGAISWEAPVEKTECIQKKNOTECFNFTRFLQPY 120
QY 121 NASHLVYCGTYAFAOPKCTYVNMFLTELEHGEFEDGKGCPCYDPAKGAHGLVDELXSAT 180
DB 121 NASHLVYCGTYAFAOPKCTYVNMFLTELEHGEFEDGKGCPCYDPAKGAHGLVDELXSAT 180
QY 181 LNLFGEPIILIRNMGPHSHSKTEYLAFWLNEMPHVSAYVPSVGSFTGDDXVYFFER 240
DB 181 LNLFGEPIILIRNMGPHSHSKTEYLAFWLNEMPHVSAYVPSVGSFTGDDXVYFFER 240
QY 241 ERAVESCYAEQVYARVAVRCGDMGARTLQKRTFLKARLACSAAPMQLYFNOLQAM 300
DB 241 ERAVESCYAEQVYARVAVRCGDMGARTLQKRTFLKARLACSAAPMQLYFNOLQAM 300
QY 301 HTLADTSNHTTFFGVFOAOWGDMYLSAICEYOLEIQRVFEQPYKEYHEBAQKMDRYTD 360
DB 301 HTLADTSNHTTFFGVFOAOWGDMYLSAICEYOLEIQRVFEQPYKEYHEBAQKMDRYTD 360
QY 361 PVSPRPGSCINNMHRNGYSSLELDPNLTNFKKRLMEQOYPRMSRPLVKKGNF 420
DB 361 PVSPRPGSCINNMHRNGYSSLELDPNLTNFKKRLMEQOYPRMSRPLVKKGNF 420
QY 421 THLVADRYTGLDGYATYVLTFTGDSGLKAVSLGPMWHLIEELQDFOEPMBSLVYSQS 480
DB 421 THLVADRYTGLDGYATYVLTFTGDSGLKAVSLGPMWHLIEELQDFOEPMBSLVYSQS 480
QY 481 KTLFAGRSQVLOLVADCTIKYSCADCVLARDPYCAMSANTSRCAV-VGCHGSLILQ 539
DB 481 KTLFAGRSQVLOLVADCTIKYSCADCVLARDPYCAMSANTSRCAV-VGCHGSLILQ 539
QY 540 HVMSTSPSGICNLNG-SKIQSGPYPKNTTVAGTDLVPLCHLSSNLA-----LPD 588
DB 540 HVMSTSPSGICNLNG-SKIQSGPYPKNTTVAGTDLVPLCHLSSNLA-----LPD 588
QY 589 SNP 591
DB 589 SNP 591

```

```

RESULT 12
US-10-189-123-65
; Sequence 65, Application US/10189123
; Publication No. US20030082586A1
; GENERAL INFORMATION:
; APPLICANT: KIRST, Susan J.
; APPLICANT: HOLTMAN, Douglas A.
; APPLICANT: FRASER, Christopher C.
; APPLICANT: SHARP, John D.
; APPLICANT: BARNES, Thomas S.
; TITLE OF INVENTION: ANTIBODIES HAVING DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE REFERENCE: 10147-1103
; CURRENT APPLICATION NUMBER: US/10/189,123
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 65
; LENGTH: 834
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-189-123-65

```

```

Query Match      84.5%; Score 2720; DB 15; Length 834;
Best Local Similarity 83.1%; Pred. No. 2,5e-267;
Matches 501; Conservative 41; Mismatches 47; Indels 14; Gaps 4;

```

```

QY 1 MAPHAWMLAARLMGIGAGAEVMMNLYPRKTVSSGELATVYVRRFSQTGIDFILTILTE 60
DB 1 MAPHAWMLAARLMGIGAGAEVMMNLYPRKTVSSGELATVYVRRFSQTGIDFILTILTE 60
QY 61 PTGLLYGAREALPAFSEALELOGAISWEAPVEKTECIQKKNOTECFNFTRFLQPY 120
DB 61 HSGLLYGAREALPAFSEALELOGAISWEAPVEKTECIQKKNOTECFNFTRFLQPY 120
QY 121 NASHLVYCGTYAFAOPKCTYVNMFLTELEHGEFEDGKGCPCYDPAKGAHGLVDELXSAT 180
DB 121 NASHLVYCGTYAFAOPKCTYVNMFLTELEHGEFEDGKGCPCYDPAKGAHGLVDELXSAT 180
QY 181 LNLFGEPIILIRNMGPHSHSKTEYLAFWLNEMPHVSAYVPSVGSFTGDDXVYFFER 240
DB 181 LNLFGEPIILIRNMGPHSHSKTEYLAFWLNEMPHVSAYVPSVGSFTGDDXVYFFER 240
QY 241 ERAVESCYAEQVYARVAVRCGDMGARTLQKRTFLKARLACSAAPMQLYFNOLQAM 300
DB 241 ERAVESCYAEQVYARVAVRCGDMGARTLQKRTFLKARLACSAAPMQLYFNOLQAM 300
QY 301 HTLADTSNHTTFFGVFOAOWGDMYLSAICEYOLEIQRVFEQPYKEYHEBAQKMDRYTD 360
DB 301 HTLADTSNHTTFFGVFOAOWGDMYLSAICEYOLEIQRVFEQPYKEYHEBAQKMDRYTD 360
QY 361 PVSPRPGSCINNMHRNGYSSLELDPNLTNFKKRLMEQOYPRMSRPLVKKGNF 420
DB 361 PVSPRPGSCINNMHRNGYSSLELDPNLTNFKKRLMEQOYPRMSRPLVKKGNF 420
QY 421 THLVADRYTGLDGYATYVLTFTGDSGLKAVSLGPMWHLIEELQDFOEPMBSLVYSQS 480
DB 421 THLVADRYTGLDGYATYVLTFTGDSGLKAVSLGPMWHLIEELQDFOEPMBSLVYSQS 480
QY 481 KTLFAGRSQVLOLVADCTIKYSCADCVLARDPYCAMSANTSRCAV-VGCHGSLILQ 539
DB 481 KTLFAGRSQVLOLVADCTIKYSCADCVLARDPYCAMSANTSRCAV-VGCHGSLILQ 539
QY 540 HVMSTSPSGICNLNG-SKIQSGPYPKNTTVAGTDLVPLCHLSSNLA-----LPD 588
DB 540 HVMSTSPSGICNLNG-SKIQSGPYPKNTTVAGTDLVPLCHLSSNLA-----LPD 588
QY 589 SNP 591
DB 589 SNP 591

```

Db 599 EOP 601

RESULT 13

US-09-989-920-218

Sequence 218, Application US/09989920

Patent No. US2002012957A1

GENERAL INFORMATION:

APPLICANT: Macina, Roberto

APPLICANT: Recipon, Heirve

APPLICANT: Chen, Sei-Yu

APPLICANT: Sun, Yongming

APPLICANT: Liu, Chenghua

TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Pro

FILE REFERENCE: DEX-0291

CURRENT APPLICATION NUMBER: US/09/989,920

CURRENT FILING DATE: 2001-11-21

PRIOR APPLICATION NUMBER: 60/252,500

PRIOR FILING DATE: 2000-11-22

NUMBER OF SEQ ID NOS: 284

SOFTWARE: Patent version 3.1

SEQ ID NO 218

LENGTH: 785

TYPE: PRT

ORGANISM: Homo sapien

US-09-989-920-218

Query Match 81.9%; Score 2635; DB 10; Length 785;

Best Local Similarity 90.0%; Pred. No. 1e-258;

Matches 495; Conservative 2; Mismatches 7; Indels 46; Gaps 4;

QY 87 ISWEAIVEKEKTECIOKKN-----OQECN 112

DB 4 ISWEAIVEKEKTECIOKKNNOYANTLLYLPSPODVSSHSPRALNTRPQTECPN 63

QY 113 FIRLOPYASHLYCGTAFOPKCTYVNLFTLEHGFEDGKCPDPAKGAHGLY 172

DB 64 FIRLOPYASHLYCGTAFOPKCTYVNLFTLEHGFEDGKCPDPAKGAHGLY 123

QY 173 DGELYSATLNLGTEPILRNKGRHNSKTEYLAFWLNEPFGVSAYVPEVSGSEFTGD 232

DB 124 DEELYSATLNLGTEPILRNKGRHNSKTEYLAFWLNEPFGVSAYVPEVSGSEFTGD 183

QY 233 DAVYFFRRRAVESDCYAOYVAVARVAVCKGDMGARTQIRKNTFFLAKRLCSAPNWL 292

DB 184 DAVYFFRRRAVESDCYAOYVAVARVAVCKGDMGARTQIRKNTFFLAKRLCSAPNWL 243

QY 293 YFNQLQAMHTLQDTSNHTTFFGVFOAGMDYLSAICEYOLEEIOREVEGPKYKEHNEA 352

DB 244 YFNQLQAMHTLQDTSNHTTFFGVFOAGMDYLSAICEYOLEEIOREVEGPKYKEHNEA 303

QY 353 QKMDRYTDPVSPRPGSCINNRHNGYTSLELDPDNLINFYKKHPLMEEOVGPRNSRPL 412

DB 304 QKMDRYTDPVSPRPGSCINNRHNGYTSLELDPDNLINFYKKHPLMEEOVGPRNSRPL 363

QY 413 LVKKGTNFTHLVADVTGIDGATYVYLFCTGDDGKYLKAVSLGPRVHLEELQLDQDEM 472

DB 364 LVKKGTNFTHLVADVTGIDGATYVYLFCTGDDGKYLKAVSLGPRVHLEELQLDQDEM 423

QY 473 RSLVVSQSK-KLLFAGSRSQLVOLPVADCIKRYSCADCVIADPVCAMSVNTRCAVAGC 531

DB 424 RSLVVSQSKKLLFAGSRSQLVOLPVADCIKRYSCADCVIADPVCAMSVNTRCAVAGC 483

QY 532 HEGSLILGHVMTSDTSGICNLNGSKIOSGPKPKNTVVAGTDLVLPCHLSNLA----- 585

DB 484 HEGSLILGHVMTSDTSGICNLNGSK-KVBPPTKNITVAGTDLVLPCHLSNLAHARWTF 542

QY 586 ----LPDSNP 591

DB 543 GGRDLPAEOP 552

RESULT 14

US-09-946-374-253

Sequence 253, Application US/09946374

Publication No. US20030073129A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Bolstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferreira, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gao, Wei-Qiang

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, Christopher J.

APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth J.

APPLICANT: Pan, James

APPLICANT: Phoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

FILE REFERENCE: P2830P1C1

CURRENT APPLICATION NUMBER: US/09/946,374

CURRENT FILING DATE: 2001-09-04

PRIOR APPLICATION NUMBER: 60/098716

PRIOR FILING DATE: 1998-09-01

PRIOR APPLICATION NUMBER: 60/098723

PRIOR FILING DATE: 1998-09-01

PRIOR APPLICATION NUMBER: 60/098749

PRIOR FILING DATE: 1998-09-01

PRIOR APPLICATION NUMBER: 60/098750

PRIOR FILING DATE: 1998-09-01

PRIOR APPLICATION NUMBER: 60/098803

PRIOR FILING DATE: 1998-09-02

PRIOR APPLICATION NUMBER: 60/098821

PRIOR FILING DATE: 1998-09-02

PRIOR APPLICATION NUMBER: 60/098843

PRIOR FILING DATE: 1998-09-02

PRIOR APPLICATION NUMBER: 60/099536

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: 60/099596

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: 60/099598

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: 60/099602

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: 60/099642

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: 60/099741

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099754

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099763

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099792

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099808

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099812

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099815

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099816

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/100385

PRIOR FILING DATE: 1998-09-15

PRIOR APPLICATION NUMBER: 60/100388

;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100390
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100584
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100627
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100661
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100662
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100664
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100683
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/100684
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/100710
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/100711
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/100848
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/100849
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/100919
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/100930
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/101014
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/101068
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/101071
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/101279
;; PRIOR FILING DATE: 1998-09-22
;; PRIOR APPLICATION NUMBER: 60/101471
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101472
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101474
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101475
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101476
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101477
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101479
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101738
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101741
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101743
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101915
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101916
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/102207
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102240
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102307
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102330
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102331
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102484
;; PRIOR FILING DATE: 1998-09-30

;; PRIOR APPLICATION NUMBER: 60/102487
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102570
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102571
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102684
;; PRIOR FILING DATE: 1998-10-01
;; PRIOR APPLICATION NUMBER: 60/102687
;; PRIOR FILING DATE: 1998-10-01
;; PRIOR APPLICATION NUMBER: 60/102965
;; PRIOR FILING DATE: 1998-10-02
;; PRIOR APPLICATION NUMBER: 60/103258
;; PRIOR FILING DATE: 1998-10-06
;; PRIOR APPLICATION NUMBER: 60/103314
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103315
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103328
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103395
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103396
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103401
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103449
;; PRIOR FILING DATE: 1998-10-06
;; PRIOR APPLICATION NUMBER: 60/103633
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/103678
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/103679
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/103711
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/104257
;; PRIOR FILING DATE: 1998-10-14
;; PRIOR APPLICATION NUMBER: 60/104987
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105000
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105002
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105104
;; PRIOR FILING DATE: 1998-10-21
;; PRIOR APPLICATION NUMBER: 60/105169
;; PRIOR FILING DATE: 1998-10-22
;; PRIOR APPLICATION NUMBER: 60/105266
;; PRIOR FILING DATE: 1998-10-22
;; PRIOR APPLICATION NUMBER: 60/105693
;; PRIOR FILING DATE: 1998-10-26
;; PRIOR APPLICATION NUMBER: 60/105694
;; PRIOR FILING DATE: 1998-10-26
;; PRIOR APPLICATION NUMBER: 60/105807

Query Match 38.5% Score 1240: DB 11: Length 837:

Best Local Similarity 45.4%: Pred. No. 8.8e-117: Mismatches 265; Conservative 88; Mismatches 201; Indels 30; Gaps 12:

OY 25 WNLVPRKTVSSGELATVVRPSQGTIGDPLTTLTTEPGTLTYVGAREALPAFMSBATELQ 84
DB 42 WALSPRISLPLGSEERPLRFEAEHISNYTALLSROGRTLYVGAREALPLSSLSLPL 101
OY 85 GA-----TWEAPVKEKTCICOKRANNOECNFTFLDPYNAHLYVCGYTAPOPKTYV 140
DB 102 GGEYOELTMGADARKQOCSEFKGDPODRCOMYIKILPLSGSHLFTGCTAFSPMCTYI 161
OY 141 NMLFTLEHGE-----FEDGKGCPYDPAKGAHGLTYGELYSATINFLGTEPIILNM 195
DB 162 MNEFTTLARDKGNVLEDEGRCRPFDPNFTSTALYVDGELYTGTVSSFOGNDPAISNSQ 221
OY 196 GPHHSMKTEYLAFMLNEPHFVSAVVPESVSGFTGDDDKVYFFPREAVESDCAEYVA 255

```
Db 222 SLRPT-KTSSSLMLODPAFVNASATPEISLGSLOQDDDKITFFSESTGGEFFENTIVS 280
QY 256 RVARVCQDMGARGARTQKRTTEFLKARLASCAPNMOLYFNQLOAMTLQDT--SMINTTF 313
Db 281 RIARICKGDGCGRRVLOQMTSLKNAQLCSRPDDFPFNVLQDVTLSBPQMDITLF 340
QY 314 FGVCQAM--GDMYLSAICEYOLEEYORVPECPYKYEHEBAQKMDRYDTPVSPRQSCI 371
Db 341 YGVTFSQAMRGTEGACVCFMVKDVOQRYFSGLYKXENETQOMYTVTHVPPTPRGACI 400
QY 372 NNMHRBHYTSSLELDPNLIINFYKKHPLMBEOYCPWMSPLLYKKGTNFTHLADVTGL 431
Db 401 TMSARBRKINSSLOLDRVNLNFKHFLMDGVR--SMMLSLQPARQORVAVHRVPGI 457
QY 432 DGATYVLEIGTGDMGLKAVSLGPMVHLEELQLEFDQ--EPMSLVLQSOKLLFAGSRS 490
Db 458 H-HTYOVLEFLGTGDLRLKHAVSGVGRVHIEELQITSSQOPQNLILDLHRLGLLYAASHS 516
QY 491 QLVQPLVADCIKTRSCADCVLARDPYCANSVNTSRGVAVGHRGSL----IOHVTSTPT 546
Db 517 GVVQVPMANCSLYRSGDILRLADPYCANS--GSSCKHVSLYOPOLATRPWIODIGASA 574
QY 547 SGICN----LRGSKIQSGXP--KNITVAGTDLVLPCHLSSNLA 585
Db 575 KDLCSASSVSPSFPVTGKPCBQVOFOPRTVNTLACPLLSNLA 618

RESULT 15
US-10-052-586-454
: Sequence 454, Application US/10052586
: Publication No. US20020127564A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jilan
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P3430R1C1
: CURRENT APPLICATION NUMBER: US/10/052,586
: PRIOR APPLICATION NUMBER: 60/059263
: PRIOR FILING DATE: 1997-09-18
: PRIOR APPLICATION NUMBER: 60/059266
: PRIOR FILING DATE: 1997-09-18
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/063120
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/063121
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/063486
: PRIOR FILING DATE: 1997-10-21
: PRIOR APPLICATION NUMBER: 60/063540
: PRIOR FILING DATE: 1997-10-28
: PRIOR APPLICATION NUMBER: 60/063541
: PRIOR FILING DATE: 1997-10-28
: PRIOR APPLICATION NUMBER: 60/063544
: PRIOR FILING DATE: 1997-10-28
: PRIOR APPLICATION NUMBER: 60/063564
: PRIOR FILING DATE: 1997-10-28
: PRIOR APPLICATION NUMBER: 60/063734
: PRIOR FILING DATE: 1997-10-29
: PRIOR APPLICATION NUMBER: 60/063870
: PRIOR FILING DATE: 1997-10-31

PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066466
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080337
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
```

```

; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086023
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086486
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087098
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087208
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088722
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088740
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088811
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088825
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088863
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089090
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12

```

```

; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089908

```

```

Query Match 38.5%; Score 1240; DB 14; Length 837;
Best Local Similarity 45.4%; Pred. No. 8-8e-117;
Matches 265; Conservative 88; Mismatches 201; Indels 30; Gaps 12;

```

```

QY 25 WNLVPRKTVSSGELATVVRSSQIGDIFLTLTEPLGLLYGARELAFPSHEALFQ 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 42 WALSPRISLPJGSEERPRPREFAEHISNYTALLSRDGRILYVGAREALFALSSNLSFLP 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 85 GA---ISMEAPVEKTECIQKKNNOTECPNFIREFIOPVNAHLYVCGTAFOPKCYV 140
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 102 GGEYQELMGADAEKQOCFRGKDPQNDONTIKILPLSGSHLFTGTAFFSMCTYI 161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 141 NMLFTLEHGE---FEDGRKCPYDPAKSHAGLYVDGLYSATLNNFTLEPIILRM 195
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 162 NMENFTLARDEKGNVLEDDGRCPDFPNFSTALYVDGELYTGTVSSFGQNDPAISRSQ 221
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 196 GPHSMKTEYLAFNLNEHFPVGSAYVPSSVSTGGDDKTYFFFRERAVESDCYAEQVYA 255
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 222 SLRPT-KTESSLNMLQDPAFVASAIPESLASLQDDDKIYFFSEFGQEEFEPEENTIVS 280
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 256 RVARCKGDMGCGARTIQRKTTFLKARLACGAPNMQLYFNOLQAMHTLQDT--SWHNTTF 313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 281 RIARCKDEGGERLQGRWTSFLKAOLCSRPDGFFPNVLDQVFTLSPSPQDRDRLTF 340
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 314 FGVEQAOV--GDMTISAICEYQLEBIOFVEFGPYKEHAEAKMDRYTDPVSPRPGSCI 371
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 341 YGVEFSQWHRGTTEGSAVCFPMKDVQVFGLYKEVRETOQWYTVTHVPPTPRGACI 400
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 372 NMHRRHGRTSSLEPDLILNFVYKHPLEMDQVGRMRRLYLKKTFTLVLADRYTGL 431
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 401 TNSAERRKINSLLDLPDVLNLFKLDHFLMDQVYR--SRMLLDQARTQVAVARVPGI 457
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 432 DCAFTYTLFEGTGDMLKKAVALPWLIELOLFDQ--EPMRSVLVSQSKLLEFAGSRS 490
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 458 H-HYDVLFLATGDRRLKAVSVGPRVHIIIELOLIFSSGQYVONLLDTHRGGLYASHS 516
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 491 QLVOLPVADCIKYSACDCVILARDPYCAMSVNTSRCAVAGGHFSGSL---IOHMYTSDT 546
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 517 GVOVPYMANCSLYRSCGCLLARDPYCAMSS--GSSCKRYSLYQPLATRPPIODIEGASA 574
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 547 SGICH---LNGSKIQSGPXP-KNITVYVAGDVLVPLHLSNYLA 585
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 575 KDLCSASSVSPSFVTEKPCQEOVQOPNTVNTLACLNLNTLA 618
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Search completed: August 7, 2003, 13:59:11
Job time : 56 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 13:13:42; Search time 3207 Seconds

(without alignments)
17309.455 Million cell updates/sec

Title: us-10-002-050-13

Perfect score: 2284
Sequence: 1 cggcccttcacacacccg.....tctgtgtgtatgtgtgtc 2284

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_estnum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rnd:*
26: em_gss_png:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1303.8	57.1	3648	AK090073	Mus muscu
2	832.4	36.4	977	BX370082	BX370082
3	831.8	36.4	932	BX369159	BX369159
4	688	30.1	864	BT518922	BT518922

5	651.4	28.5	929	13	BQ917819	BQ917819	AGENCOURT
6	632.6	27.7	705	10	BE294211	BE294211	AGENCOURT
7	621.2	27.2	937	13	BQ151848	BQ151848	AGENCOURT
8	603.8	26.4	1394	12	BM926974	BM926974	AGENCOURT
9	597.6	26.2	768	12	BM943605	BM943605	AGENCOURT
10	597.6	26.1	632	10	BE270240	BE270240	AGENCOURT
11	579.6	25.4	1002	13	BQ068313	BQ068313	AGENCOURT
12	575.4	25.2	853	13	BX396978	BX396978	AGENCOURT
13	569.2	24.9	745	12	BM629654	BM629654	AGENCOURT
14	554.6	24.3	763	13	BQ612234	BQ612234	AGENCOURT
15	536.4	23.5	597	10	BQ324157	BQ324157	AGENCOURT
16	536	23.5	1075	13	BX360753	BX360753	AGENCOURT
17	527.4	23.1	978	13	BQ963608	BQ963608	AGENCOURT
18	520.2	22.8	738	12	BM414643	BM414643	AGENCOURT
19	507.4	22.2	965	13	BQ445252	BQ445252	AGENCOURT
20	501.8	22.0	977	13	BQ855862	BQ855862	AGENCOURT
21	494.8	21.7	647	12	BM950633	BM950633	AGENCOURT
22	487.6	21.3	673	12	BM951984	BM951984	AGENCOURT
23	477	20.9	789	12	BM86539	BM86539	AGENCOURT
24	452	19.8	1033	13	BX460723	BX460723	AGENCOURT
25	443.4	19.4	604	10	BE283742	BE283742	AGENCOURT
26	435.4	19.1	455	14	W01259	W01259	AGENCOURT
27	435	19.0	1129	13	BQ071719	BQ071719	AGENCOURT
28	431.2	18.9	766	12	BM45497	BM45497	AGENCOURT
29	428	18.7	533	13	BQ266902	BQ266902	AGENCOURT
30	428	18.7	923	12	BI665624	BI665624	AGENCOURT
31	397.8	17.4	704	14	BY741417	BY741417	AGENCOURT
32	395.4	17.3	710	9	AI638881	AI638881	AGENCOURT
33	391.2	17.1	840	14	CB990331	CB990331	AGENCOURT
34	390.4	17.1	404	12	BM174889	BM174889	AGENCOURT
35	387.6	17.0	911	13	BQ278955	BQ278955	AGENCOURT
36	378.6	16.6	3763	11	AK088653	AK088653	Mus muscu
37	376.8	16.5	2777	11	AK028896	AK028896	Mus muscu
38	376.2	16.5	659	9	AI651838	AI651838	Mus muscu
39	375.2	16.4	3809	11	AK030331	AK030331	Mus muscu
40	372.4	16.3	707	14	BY733150	BY733150	Mus muscu
41	367	16.1	690	14	BY734086	BY734086	Mus muscu
42	358.8	15.7	573	10	BE070335	BE070335	QV4-BT040
43	343.4	15.0	453	14	CB060019	CB060019	4010021 B
44	340	14.9	680	14	BY735266	BY735266	BY735266
45	336.2	14.7	543	14	CD216128	CD216128	P9P2n.PX0

ALIGNMENTS

-RESULT 1
LOCUS AK090073
DEFINITION Mus musculus 7 days embryo nullipotent stem cell 1 CR1-2070 NE cDNA, RIKEN full-length enriched library, clone:G430091K07 product:soma domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4C, full insert sequence.
ACCESSION AK090073
VERSION AK090073.1 GI:26105718
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374

PUBMED REFERENCE AUTHORS	11042159 3 Shibata,K., Itoh,M., Aizawa,K., Nagecke,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,Y., Nishik,K., Katsuna,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,N., Nishie,T., Harada,A., Yamamoto,R., Matsuno,H., Saezangchi,S., Ikegami,T., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN Integrated sequence analysis (RISA) system -384-format Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	20530913 4 11076861 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
TITLE JOURNAL REFERENCE AUTHORS	5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
TITLE JOURNAL REFERENCE AUTHORS	6 (bases 1 to 3648) Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Hori,F., Imocani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishik,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Salto,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE JOURNAL	Direct Submission Submitted (16-APR-2002) Yoshinori Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-2 Suhei-cho, Tsukuba-shi, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-reg@gs.riken.go.jp, URL: http://genome-gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome-gsc.riken.go.jp/ URL: http://fantom-gsc.riken.go.jp/.
FEATURES SOURCE	Location/Qualifiers 1..3648 /organism="Mus musculus" /mol_type="mRNA" /strain="C3H/3" /db_xref="FANTOM.DB:G430091K07" /db_xref="taxon:10090" /clone="G430091K07" /cell_line="CRL-2070 NE" /cell_type="nullipotent stem cell" /clone_id="RIKEN full-length enriched mouse cDNA library /dev_stage="7 days embryo" 245..2752 /note="putative zeta domain, immunoglobulin domain (Iz), transmembrane domain (TM), and short cytoplasmic domain, (semaphorin) 4C (MGI:MGI:109252, GBIS79463, evidence: BLASTN, 98%, match=3433)"
misc_feature	
BASE COUNT	728 a 985 c 1089 g 844 t 2 others

Query Match	57.1%	Score 1303.8	DB 11	Length 3648
Best Local Similarity	83.4%	Pred. No. 3.2e-298		
Matches 1542	Conservative 0	Mismatches 296	Indels 10	Gaps 5
OY	97	GGACCTCGCGTCGCGAGCCGGCGCCACACGAGGAGACAGGCGCTGACCTGAGAGCCGG	156	
DB	176	GCACCAACCGGGTTTGGACCCCAAGCCCTAGCGACAGGGCTGGGCTGACATTAGACCCTGG	235	
OY	157	GCCTGCGCCATGAGCCCACTAGGCGCTGTGGCTGTGGCTGTGGCAGCAAGGCTGTGGGCGTG	216	
DB	226	GTGTCTGCATATGGCCCCACACACTGGGCTGTGGGCTGTGGGCTGTGGGAGAGGCGTGTGGCTGTG	295	
OY	217	GGCATTTGGGCTGAGAGTGTGGTGAACCTTTGTGCCGCTTAAGACAGTGTCTTGTGGGAG	276	
DB	296	GGCATCTGGGGGTGAGATGTGGTGAACCTTTGTGCCGCTTAAGACAGTGTCTTGTGGGAG	355	
OY	277	CTGGCCACAGGATGATGCGGGGCTTCCCAAGCCGGCATCCAGAGATCTCCATGCACTGACG	336	
DB	356	CTGGCTACAGATGATGAGCGGGTCTCCCAAGACAGGCAATCCAGAGATCTCCATGCACTGACG	415	
OY	337	CTGACGAGACCCACTAGGGCTCTGTACGTGGGGCGCCGAGAGAGGCGCTTTTGGCTTCAACG	396	
DB	416	CTGACAGAAACATTCGGGCTTTTATATGTGGGGCGCCGAGAGGCGCTTTTGGCTTCAACG	475	
OY	397	ATGGAGCGCTGTGAGCTGCAAGGAGCGATCTTCCTGGGAGGCGCCCGTGGAGAAAGAACT	456	
DB	476	GTAGAGCGCTGTGAGCTGCAAGGAGCGATCTTCCTGGGAGGCGCTTCAGTACAGAAATTT	535	
OY	457	GAGGTATTCAGAAAAGGGAAGAACAAACGAGCCGATCTTCAATCTATCCGCTTCTGTG	516	
DB	536	GAAATATACCCAGAAAAGGGAAGAACAAACGAGCCGATCTTCAATCTATCCGCTTCTGTG	595	
OY	517	CAGCCCTTACAA - TGCCCTCCCACTGTATAGTGTGGACCTTACGCGCTTCCAGCCCAATGTG	575	
DB	596	CAGCCATTAACCTTCTCCCAATGTGTATGTGTGGGACACCTATGCTTCCAGCCCAATGTG	655	
OY	576	CACCTACGTCAACATGCTCACCTTCACCTTTGGAGCATGAGAGCTTGAAGATGGAGAGG	635	
DB	656	CACCTACATCAACATGCTCACCTTCACCTTTGGAGCATGAGAGCTTGAAGATGGAGAGG	715	
OY	636	CAAGTTCGCCATATGACCCAGCTTAAGGGGCAATGTGCGCTTCTGTGGATATGGAGACTGTAT	695	
DB	716	TAAATGCCCATATGACCCAGCTTAAGGGGCTACACCGGAGCTCTCTGTGGAGGGTACCTGTAT	775	
OY	696	CTCGGCCACACTCAACAACCTTCTGTGGGACAGGAAACCAATTAATCTGCTTAACATGGGGCC	755	
DB	776	CTCAGGCCACACTCAATACTTCTGTGGGACAGAGCGGCTTATCTTCATATACATGGGGAC	835	
OY	756	CCACCACTCATGAAGACAGAGTACCTGGCTTTTGGCTCAACGAACTCACTTTGTATAGG	815	
DB	836	CCACCACTCATGAAGACAGAGTACCTGGCTTTTGGCTCAATGAATGAACCCACTTTGTATAGG	895	
OY	816	CTCTGCCATATATACCTGAGAGTGTGGGACGCTTCAAGGGGGAGACAGCAAGGTACTT	875	
DB	896	CTCTGCCATATATACCTGAGAGTGTGGGACGCTTCAAGGGGGAGACAGTCAAGATCTACTT	955	
OY	876	CTTCTTCAGGAGCGGCGAGTGGATGCCGATATCCGAGCAGGATGGGCTCGTGT	935	
DB	956	CTTCTTCAGTACGCGGCGAGTGGATGATGACTGCTATATCCAGACAGGATGGGCTCGTGT	1015	
OY	936	GGCCCGTGTCTGCAAGGGCGATATGGGGGGCGCAGGAGCCCTGTGAGAGAGTGGACACG	995	
DB	1016	GGCGAGAGTCTGTAAAGGTGACATGAGGGGGAGACAGGAGCGCTGAAGAAATGAGAGAC	1075	
OY	996	GTTCCCTGAAGCGCGGCTGGCATGTCTGTGCCCGAATGGCAGCTGTACTTCAACACAGCT	1055	
DB	1076	GTTCTCTGAAGCGCTGGTGTGTGTGCTCAGCCCTGTACCTGAGGAAGGCTCTACTTCAACACAGCT	1135	
OY	1056	GCAGGCGATGATGACACCCGACAGAACACTCTCTGGGACAAACACACCTCTTTGGGGTTTT	1115	
DB	1136	GAAAGCGGATGACACCCGCGGGGACCTCTTGGGACAAACACACCTTCTTCCGGGGTTTT	1195	

OY	1116	TCAAAGCAGAGGGGGTGCATGTCACCTGCTGGGCCATCTGTAGACAGAGTTGGAAAGAT	1175
Db	1196	TCAAGCCGGATGGGGCATATGAGACCTGTCTGCAGTTTGTAGTACCACTTTGGACAGAT	1255
OY	1176	CCAGCGGGGTGTTGGAGGCCCTTATTAAGAGTACCATGAGGAAAG--CCCAGAAAGTGGGAC	1233
Db	1256	CCACCAAGTGTTTTAGGGGTCCCTCAAGAGAGTACAGTACGACAGAGGCCAGAAAGTGGGGCC	1315
OY	1234	CGCTACACTGACCCCTGATCCAGGCCCTCGGCTGGCTGTGCATTTAAACAATTGGCATTCGG	1293
Db	1316	CGCTATATGACCCGGTATCCAGCCCTCGGCTGGTGTGTATCAACAATGTGCACCGA	1375
OY	1294	CGCCACGGCTACACAGCTCCCTGG--AGCTACCCGACAAACATCTCTCAACTCTGTCAGAA	1352
Db	1376	GACAAATGGGTACACAGTGTCCCTGGAAAGCTCCGGACACACACCTCAACTCATCAAGAA	1435
OY	1353	GCACCCGGTATGGAGGAGCAGTGGGGCCTCGGTGAGCGCCGCCCTCTCTGTGAAGAA	1412
Db	1436	GCACCCCTGTATGGAGGACCAAGGTGAAGCCCTGGGTGGCGCCGCCCTCACTTGTGAAGAA	1495
OY	1413	GGGACCAACCTTCACCCCACTGGTGGGCCGACGGGGTTACAGGACTTGAATGAGACCACTA	1472
Db	1496	GAACTACTACTTCACACAGCTGGTGGCCGAGAGGATCCACAGGGTGTGATGGTGCACACTA	1555
OY	1473	TACAGTCTGTTCATTGTGGACAGAGAGAGCGCTGCTCTCAAGGCTGTGACCTGGGGCC	1532
Db	1556	TACAGTGTGTTCATGTGTACAGCAGAGATGGCTGGCTGCTGAAGGCTGTGAGACTGGGGCC	1615
OY	1533	CTGGGTTCACTGATTTGAGAGCTGACAGCTGTTTGAACAGGAGCCCATGAGAAAGCTGTG	1592
Db	1616	CTGGATCCACATGTGGAGGAACTGACAGGTGTTTGAACAGGAGCCAGTGGAAATCTGTGT	1675
OY	1593	GCTATCTAGAGCAAGAGCTGCTCTTGGCGGCTCCCGCTCATAGCTGGGAGAGCTGCC	1652
Db	1676	GCTGCTCTAGACAGAAAGTGTCTTTGTGGGCTCCCGCTCTCAAGCTGGGTTCAGCTGTC	1735
OY	1653	CGTGGCCGACTGCTAAAGTATCGCTCTCTGTGCAGACTGTGTCTCTGCGCCGGAGACCTTA	1712
Db	1736	TCTGGCCGACTGACAAAGTACCGTTTGTGTAAACTGTGTCTCTGGCCAGGAGACCTTAA	1795
OY	1713	TTGGGCGCTGAGAGCTCAACACACAGCGCTGTGTGTG--CCGTGGGTGGCCACTTTGGATC	1769
Db	1796	CTGTGCTCGGATGTTCGACACACACCGCTGTGTGTGGCCACCAACAGTGTCTCGGGGCTC	1855
OY	1770	TTTACATGATTCAGCACTGTGATGACCTCGGACACTTCAAGCATTTTGCAACTCTCCGGGAG	1829
Db	1856	CTTTCTGGTCCAAACATGTGGCGAACTTGGACACTTCAAAAGATGTGTAAACCAAGTATGGCAT	1915
OY	1830	TAAGATACAGTCAGGCCCACTNCCCAAAAACATCAAGGTGTGTGGCGGGCAGACAGACTGTG	1889
Db	1916	TAA--AAAAGTACAGATCTATTATCCCAAGAACATCAACCTGTGTGTACGAGCACAGACCTGT	1972
OY	1890	GCTGGCCCTGCACCACTCTCTCCAACTTGGCCCTGCGCGACTCCAAACCCG	1939
Db	1973	CTTACCTGGCCACCTCTCTCCAAATTTGGGCCCAATGCCACCTGAGACCTTTCG	2022

[illegible]

JOURNAL
COMMENT

Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seque@genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1123.r. For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BA6036ZC10_CS03435_1&cluster=1123.r.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BA6036ZC10_CS03435_1.

FEATURES
source

1..977
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1074YD21"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_1lb="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT
204 a 289 c 277 g 207 t

ORIGIN

Query Match 36.4%; Score 832.4; DB 13; Length 977;
Best Local Similarity 93.9%; Pred. No. 1.6e-181;
Matches 900; Conservative 0; Mismatches 51; Indels 7; Gaps 3;

OY 496 TTCAACTTCATCCGCTTCCTGACAGCCCTACAAATGCTCCACCTGACGTACTGTGGAC 555
DB 1 TTCAACTTCATCCGCTTCCTGACAGCCCTACAAATGCTCCACCTGACGTACTGTGGAC 60
OY 556 TACGCTTCACGCCCAAGTGCACCTAGTCAACATGCTCACCTTCACCTTTGGACATGA 615
DB 61 TACGCTTCACGCCCAAGTGCACCTAGTCAACATGCTCACCTTCACCTTTGGACATGA 120
OY 616 GAGTTTGAAATGGGAAGGCAAGTGTGTCCTATGACCAGCTAAGGGGCAATGTGGCCTT 675
DB 121 GAGTTTGAAATGGGAAGGCAAGTGTGTCCTATGACCAGCTAAGGGGCAATGTGGCCTT 180
OY 676 CTGTGTGATGTGAGCTGTACTCGGCCACACTCAACAACTTCCTGGGCACGAAGCCATT 735
DB 181 CTGTGTGATGTGAGCTGTACTCGGCCACACTCAACAACTTCCTGGGCACGAAGCCATT 240
OY 736 ATCTCGCTAACATATGGGGCCCCACACACTCATGAAGACAGTACCTGGCTTTGGCTC 795
DB 241 ATCTCGCTAACATATGGGGCCCCACACACTCATGAAGACAGTACCTGGCTTTGGCTC 300
OY 796 AAGCAACCTCACTTTAGTGTGCGCTATGATCTAGTGAATGTGGGACGTTACAGGGG 855
DB 301 AAGCAACCTCACTTTAGTGTGCGCTATGATCTAGTGAATGTGGGACGTTACAGGGG 360
OY 856 GACGACGACAAAGTCTAATCTTCTTCAGGAGCGGCGATGAGTCCGATCGTATATCC 915
DB 361 GACGACGACAAAGTCTAATCTTCTTCAGGAGCGGCGATGAGTCCGATCGTATATCC 420
OY 916 GAGCAGTGTGTGCTCGTGTGGCCGCTGTGCAAGGGCGATATGGGGGGGCGACGGAC 975
DB 421 GAGCAGTGTGTGCTCGTGTGGCCGCTGTGCAAGGGCGATATGGGGGGGCGACGGAC 480
OY 976 CTGCAGAGGAAGTGGACACTTCTCTGAAGCGCGGCTGGCATCTCTGCCCCGAACTGG 1035
DB 481 CTGCAGAGGAAGTGGACACTTCTCTGAAGCGCGGCTGGCATCTCTGCCCCGAACTGG 540
OY 1036 CAGCTCTACTTCAACACAGCTCAGGCGCATGACACACCTGACAGACACTCTGGCACAAAC 1095
DB 541 CAGCTCTACTTCAACACAGCTCAGGCGCATGACACACCTGACAGACACTCTGGCACAAAC 600
OY 1096 ACCACTTCTTTGGGGTTTTCAAGCACACAGTGGGGTGAATGATCACTGTGGCCACTGT 1155

Db 601 ACCACCTCTTGGGTTTAAAGCAGAGTGGGTGACATCTACCTCTGGCCATCTGT 660
1156 GAGTACAGATTGGAAGATCCAGCGGGTGTGTGAGGGCCCCCTATAGAGATACATGAG 1215
661 GAGTACAGATTGGAAGATCCAGCGGGTGTGTGAGGGCCCCCTATAGAGATACATGAG 720
1216 GAGGCCGGAAGTGGAGCGGCTACACTGACCTGTGCCACCTTGGGCTGGCTGGCC 1275
721 GAGGCCGGAAGTGGAGCGGCTACACTGACCTGTGCCACCTTGGGCTGGCTGGCC 780
1276 ATTACACATGGGATCGGGCGGCGGCTACACCCAGCTCCCTGGAGCTACCCGACACATC 1335
781 AATACACATGGGATCGGGCGGCGGCTACACCCAGCTCTTGGAGCTACCGACACATC 840
1336 CTCACATTCGTTCAGAGACACCCGCTGATGAGAGAGTGGGGCTCTGGTGGAGCCG 1395
841 CTAAATTTGTTAGAA---CAACCGGTATGGAGAGAGAGTGGGGCTCTGGTGGAGCC 897
1396 CCCCTGCTCTGAGAGAGGCGACCACTTCAACCCAGCTGTGGCCGACCGGGTTACAG 1453
898 GCCCTGCTCGGGA--AAAGGCACCAACTT--ACCCCTGGGGGCGCCGCTGCACGG 951

RESULT 3

LOCUS BX369159 932 bp mRNA linear EST 08-MAY-2003

DEFINITION BX369159 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0D1027YN09 5-PRIME, mRNA sequence.

ACCESSION BX369159
VERSION BX369159.1 GI:30455790

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 932)

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT Contact: genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 1123.r for

more information about this cluster, see

http://www.genoscope.cns.fr/

cgl-bln/cluster/cgl1seq-CS0BAF0242B08.AF02267.26cluster-1123.r.

Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/Invitrogen corporation 1600

Paradise Avenue Genoscope sequence ID: CS0BAF0242B08.AF02267_2.

FEATURES

source

1..932

Location/Qualifiers

1..932

1..932

1..932

Db 1 CTTCAGGCCCAAGTGCACCTACCTAAGATGCTACCTTATTGGAGCATGAGACTT 60
621 TGAAGATGGGAAGGCAAGTGTCCCTATGACCAGCTAAGGCGCATGCTGCTTGT 680
61 TGAAGATGGGAAGGCAAGTGTCCCTATGACCAGCTAAGGCGCATGCTGCTTGT 120
681 GATGATGAGCTGTACTGGGCGACACTCAACACTTCTGGGCGAGGAACCATTAATCT 740
121 GATGATGAGCTGTACTGGGCGACACTCAACACTTCTGGGCGAGGAACCATTAATCT 180
741 GGTAAACATGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 800
181 GGTAAACATGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
801 ACCCTCACTTTGAGCTCTGCTTATGACCTGAGAGTGTGGGCGACTTACCGGGGAGCA 860
241 ACCCTCACTTTGAGCTCTGCTTATGACCTGAGAGTGTGGGCGACTTACCGGGGAGCA 300
861 CGACAAAGGTCTACTTCTTTCAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 920
301 CGACAAAGGTCTACTTCTTTCAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 360
921 GGTGATGCTGTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 980
361 GGTGATGCTGTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
981 GAGGAAGTGGAGCACAGCTTCTGTAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1040
421 GAGGAAGTGGAGCACAGCTTCTGTAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 480
1041 CTACTTCAACAGCTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1100
481 CTACTTCAACAGCTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
1101 CTCTTTGGGGGTTTTCAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1160
541 CTCTTTGGGGGTTTTCAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
1161 CCAGTTGGGAAGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1220
601 CCAGTTGGGAAGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
1221 CCAGGAAGTGGAGCACAGCTTCTGTAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1280
661 CCAGGAAGTGGAGCACAGCTTCTGTAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
1281 CAACTGGCATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1340
721 CAACTGGCATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 779
1341 CTTCGTCAAGAAAGCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1398
779 CTTCGTCAAGAAAGCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 839
1399 CTTCGTCAAGAAAGCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1457
839 CTTCGTCAAGAAAGCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 899
1458 TGATGAGCGCACCTATACAGCTGCTTCAATG 1489
899 TGATGAGCGCACCTATACAGCTGCTTCAATG 931

RESULT 4

LOCUS B1518922 864 bp mRNA linear EST 29-AUG-2001

DEFINITION B1518922 B1518922 Homo sapiens cDNA clone IMAGE:5210938 5', mRNA sequence.

ACCESSION B1518922
VERSION B1518922.1 GI:15343714

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

1..864

1..864

1..864

1..864

1..864

1..864

1..864

1..864

1..864

1..864

1..864

1..864

1..864

1..864

1..864

1..864

1..864

1..864

1..864

1..864

1..864

1..864

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 864)	NIH-MGC	http://mgc.nci.nih.gov/	Unpublished
2	National Institutes of Health, Mammalian Gene Collection (MGC)	Contact: Robert Strausberg, Ph.D.	Email: cgs@bbs-remail.nih.gov	Tissue Procurement: Life Technologies, Inc.
3	CDNA Library Preparation: Life Technologies, Inc.	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	DNA sequencing by: Incyte Genomics, Inc.	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
4	Plate: LHAM1529	row: h	column: 11	High quality sequence stop: 818.
5	Location/Qualifiers	1..864		
6	/organism="Homo sapiens"	/mol_type="mRNA"	/db_xref="taxon:9606"	/clone="IMAGE:5210938"
7	/tissue_type="leukocyte"	/lab_host="DH10B"	/clone_lib="NIH-MGC_118"	/note="Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH-MGC Library."
8	BASE COUNT	167 a	244 c	296 g
9	ORIGIN	157 t		
10	Query Match	30.1%	Score 688;	DB 12; Length 864;
11	Best Local Similarity	93.9%;	Pred. No. 3.6e-148;	
12	Matches 738; Conservative	0;	Mismatches 45;	Indels 3; Gaps 2
13	84	CAGGCCGGGACCGGAGACTCCGCTGGCAGCCGCGCACACGACGACAGCTGGCTGAA	143	
14	Db	79	CCGAACCGGAGGGCGCGCCGCGGAGCTGGAGCTCGCGCAATAGACAGCTGGCTGAA	138
15	Oy	144	GCTCAGAGCCGGGGG-CGTGGCGCATAGGCGCCACACTGGGGTGTGGCTGTGGCAGCA	202
16	Db	139	GCTCAGAGCCGGGGCTGTGGCCCATAGCCCCACACTGGGCTGTGGCTGTGGCAGCA	198
17	Oy	203	GCGTGTGGGCGCTGGGCAATGGGGCTGAGTGTGTGTGGCAACCTTGTGCCGCTAAGACAG	262
18	Db	199	GGCTGTGGGCGCTGGGCAATGGGGCTGAGTGTGTGTGGCAACCTTGTGCCGCTAAGACAG	258
19	Oy	263	TGTCTTCTGGGGAGCTGGCCACAGGTAGTACGGCGGTTCTCCACAGCCGACATCAGACT	322
20	Db	259	TGTCTTCTGGGGAGCTGGCCACAGGTAGTACGGCGGTTCTCCACAGCCGACATCAGACT	318
21	Oy	323	TCTTGACACTGACGCTGACGGAGGCCACATGGGCTTCTAGCTGGGGCGCCGACAGAGCC	382
22	Db	319	TCTTGACACTGACGCTGACGGAGGCCACATGGGCTTCTAGCTGGGGCGCCGACAGAGCC	378
23	Oy	383	TGTTTGCCCTTACGATGGAGGCCCTGGAGACTTGCAGAGACGATCTCTGGAGAGGCCCG	442
24	Db	379	TGTTTGCCCTTACGATGGAGGCCCTGGAGACTTGCAGAGACGATCTCTGGAGAGGCCCG	438
25	Oy	443	TGAGAGAGAGACTGAGTGTATCCAGAAAGGAGAGAACACGACCGAGTCTTCAACT	502
26	Db	439	TGAGAGAGAGACTGAGTGTATCCAGAAAGGAGAGAACACGACCGAGTCTTCAACT	498
27	Oy	503	TGATCCGCTTCTGTGACGCTTACATATGCTCTCCACTGTAGCTGTGTGGCACTTACGCT	562
28	Db	499	TGATCCGCTTCTGTGACGCTTACATATGCTCTCCACTGTAGCTGTGTGGCACTTACGCT	558

OY		563	TCACGGCCCAAGTGCACCTAGCTGTAACAATGCGACACTCCTTGAGTGGACATGGAGAGTTG	622
Db		559	TTCAGGCCCAAGTGCACCTAGCTGTAACAATGCTCACCTTGAGTGGACATGGAGAGTTTG	618
OY		623	AAGATGGGAAGGGCAAGTGTCCCTCATGACCAGCTAAGGGCCATGTCGGCCTTCTTGG	682
Db		619	AAGATGGGAAGGGCAAGTGTCCCTCATGACCAGCTAAGGGCCATGTCGGCCTTCTTGG	678
OY		683	ATGTGAGCTGTACTCGGCGCACACTCAACAACCTCTGGGCGAGGAACCATATATCTGC	742
Db		679	ATGTGAGCTGTACTCGGCGCACACTCAACAACCTCTGGGCGAGGAACCATATATCTGC	738
OY		743	GTAACATGGGGCCCCACACTCCATCCATGAAAGAGAGTACCTGGCTTTGGCTCAACGAC	802
Db		739	GTAACATGGGGCCCCACACTCCATCCATGAAAGAGAGTACCTGGCTTTGGCTCAACGAAA	798
OY		803	CTCATTCTGTAGG--CTCTGCTCATGTATACCTGAGAGTGTGGGCGAGCTTCACGGGGGAGCA	860
Db		799	CTCATTCTGTAGGCGCTGTGGCTATGTATACCTGAGAGTGTGGGCGAGGTTCACCGGGGGGAC	858
OY		861	CGACAA 866 	
Db		859	CGACGA 864	
RESULT 5				
LOCUS	B0917819			
DEFINITION	AGENCOURT_8752723 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6335699	929 bp	mRNA	linear EST 20-AUG-2002
ACCESSION	B0917819			
VERSION	B0917819.1	GI:22332517		
KEYWORDS	EST.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	NIH-MGC http://mgc.ncl.nih.gov/.			
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs@email.nih.gov Tissue Procurement: Mark Macconochie, Ph.D. and Nancy L. Freeman, Ph.D. cDNA Library Preparation: Resgen, Invitrogen Corp CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution Information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.nih.gov Plate: LLM13798 row: 1 column: 12 High quality sequence stop: 704. Location/Qualifiers 1..929 /organism="Mus musculus" /mol_type="mRNA" /db_xref="taxon:10090" /clone="IMAGE:6335699" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_130" /note="Organ: otcyasts; Vector: pCMV-Sport6.1.ccdp; Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.95 kb. Constructed by Resgen, Invitrogen Corp. Note: this is a NIH_MGC library."			
FEATURES	source			
BASE COUNT	193 a	242 c	285 g	207 t 2 others
ORIGIN				
Query Match	28.5%	Score 651.4;	DB 13;	Length 929;
Best Local Similarity	84.5%;	Prod. No. 1e-139;		
Matches 756;	Conservative 0;	Mismatches 133;	Indels 6;	Gaps 2;

	Accession	Source	Organism
OY	823	TATGACCTAGAGAGTGTGGGACGCTTACGGGGGAGCAGCAGCAAGCTCTACTCTTCTTC	882
OY	1	TTTGTCCCTGAGAGTGTGGGAAGCTTACCGGGAGCAGTACAGATCTACTTCTTCTTC	60
OY	883	AGGAGCGGGGAGTGGAGTCCGACTGCTATGCGGAGCAGTGGTGGCTGTGTGGCCGT	942
Db	61	AGTGAAGCGGGAGTGGAGTATGACTGCTATTCCGAGCAGTGGTGGCTGTGTGGCGAGA	120
OY	943	GTCCTGCAAGGCGGATATATGGGGGGCCACAGAACCTTCGACAGAGTGGACACGTTCTCTG	1002
Db	121	GTCT-----GGTGACATGGGGGGAGACAGACGCTGCAGAAAGAAATGAGCAGCGTTCTGTG	175
OY	1003	AAGCGCGGCGCTGGCATCTCTGCGCCGAGACGAGCTCACTACTTCAACAGCTGCAGGCG	1067
Db	176	AAGCTGCGGTTGTGTGTCTCAAGCCCTTACCTGGAAGGTCTACTTCAACAGCTGAAGGCG	235
OY	1063	ATGCACACCCCTGCAGAGACACTCTGTGGCAACACACACTCTTCTTGGGGTTTTTCAMCA	1122
Db	236	GTGCACACCCCTGCGGGGCGCCCTTGTGGCAACACACACTCTCTTCGGGGTTTTTCAMCG	295
OY	1123	CAGTGGGGTGACATGTATCTGTCCGGCCATCTGTGATGCAGATTGGAGAGATCCAGGG	1187
Db	296	CGATGGGGCGATATGTAGACTGTCTGCACATTTGTGTGATGCAGATTGGAGAGATCCAGGA	355
OY	1183	GTGTTTGAAGGCGCCCTATATGAAGATACATATGAAGGCGCAAGTGGAGCGACTACT	1247
Db	356	GTGTTTGAAGGCTCCCTACAGGAATACAGTGAAGCAAGCCAGAAATGGGCCCGCTATCT	415
OY	1243	GACCTGTATCCAGACGCTCTGGCGCTGTGTGTGATTAACACTGGCATTCGGCGCCAGGCG	1307
Db	416	GACCCGGATGCCACAGCCCTCTGGCGCTGTGTATCAACAACTGGCACCGAGACATATGGC	475
OY	1303	TACACACACTCTCCCTGGAGCTACCCGACAAATCCCAACTTCGTCAAGACACCCGGTG	1362
Db	476	TACACACACTCTCCCTGGAGCTACCCGACAAACCTTCAACTTCATCAAGACACCCCTG	535
OY	1363	ATGAGAGAGCAGTGGGGGCTCTGGGTGAGCCGCCCTCTCTGTGTAAGAAAGGACACAAAC	1422
Db	536	ATGAGAGAGCAGTGGAGCCTCTGGGTGGGGCGCCCTCACTTGTGAAGAAAGACACTAAC	595
OY	1423	TTCAACCCACCTGGGTGGCCGACCGGCTTACAGAGCTTGTATGAGAGCACTATACAGTGTG	1488
Db	596	TTCAACACAGTGGGTGGCCGACAGGCTTGTATGTGTGCACACTAATACAGTGTGTG	655
OY	1483	TTCAATTGGCAGAGAGAGACGGCTGGGTGTCTCAAGGCTGTGAGACCTGAGGGGCCCTGGGTTGAC	1542
Db	656	TTCAATTGGTACAGAGAGATGGCTGGCTGTGAAGGCTGTGAGACCTGAGGGGCCCTGGATCCAC	715
OY	1543	CTGATTGAGAGCTGCAGCTGTTTGACCAAGAGCCCATGAGAAAGCTGTGTATCTACG	1602
Db	716	ATGATGGAGAGAACTGCAGAGTGTGTTGACCAAGAGCCATGGAAGTCTGTGTCTGTCTCAG	775
OY	1603	AGCAAGAAAGCTGTCTTTTGGCCGCTCCCGCTCTCAAGCTGTGTGAGAGTCCCGTGGCCAGC	1662
Db	776	AGCAAGAAAGGTGTCTTTTGTGTGTCTCCCGCTCTCAAGCTGTGTGAGAGTCTCTGTGGCCAGC	835
OY	1663	TGCATTAAGATATCG-CTCCTCTGCAAGACTGTGTCTGCGCGGAGACCCCTATATGC	1716
Db	836	TGCACAAAGTACCGTTTCTGTGTATGACTGTGTCTGTGGNKAAGGGAGCCCTTTAC	890
RESULT 6	BE294211	705 bp	mRNA linear EST 20-JUL-2000
LOCUS	BE294211		
DEFINITION	60117274AF1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:3528227 5',		
ACCESSION	BE294211		
VERSION	BE294211.1	GI:9177845	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 705)				
NIH-MGC	http://mgc.ncl.nih.gov/.			
National Institutes of Health, Mammalian Gene Collection (MGC)				
Unpublished				
Contact: Robert Strausberg, Ph.D.				
Email: csapbs@remail.nih.gov				
Tissue Procurement: ATCC				
cDNA Library Preparation: Ling Hong/Rubin Laboratory				
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)				
DNA sequencing by: Incyte Genomics, Inc.				
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: image.lnl.gov				
Plate: L1CMI197	row: 9	column: 12		
High quality sequence stop: 624.				
Location/Qualifiers				
1..705				
/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="taxon:9606"				
/clone="IMAGE:3528227"				
/tissue_type="rhabdomyosarcoma"				
/lab_host="DH10B (phage-resistant)"				
/clone_lib="NIH_MGC_17"				
/note="Organ: muscle; Vector: pOT57; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."				
BASE COUNT	143 a	209 c	215 g	138 t
ORIGIN				
Query Match	27.7%	Score 632.6;	DB 10;	Length 705;
Best Local Similarity	98.2%	Pred. No.2.1e-135;		
Matches	661;	Conservative	0;	Mismatches 9;
			Indels	3;
			Gaps	2;
749	TGGGGCCCGCACCTCCATCATGAAGACAGAGTACCTGGCCCTTTTGGCTCAGCAAGACCTCACT	808		
Db	1	TGGGGCCCGCACCTCCATCATGAAGACAGAGTACCTGGCCCTTTTGGCTCAGCAAGACCTCACT	60	
Oy	809	TTGTAAAGGCTCGCCTATGTACTTGAAGATGTGGGACAGTTTACGGGGGAGCAGCACAAG	868	
Db	61	TTGTAAAGGCTCGCCTATGTACTTGAAGATGTGGGACAGTTTACGGGGGAGCAGCACAAG	120	
Oy	869	TCTACTTCTTTTCAGAGGAGCGGGCAGTGAAGTCCGACTGCTATGCCAGCAGAGGTGGTG	928	
Db	121	TCTACTTCTTTTCAGAGGAGCGGGCAGTGAAGTCCGACTGCTATGCCAGCAGAGGTGGTG	180	
Oy	929	CTCGTGTGCCCCGTGTCTGCAAGGGCGCATATGGGGGGCGCACGGACCCCTGCAGAGAAAT	988	
Db	181	CTCGTGTGCCCCGTGTCTGCAAGGGCGCATATGGGGGGCGCACGGACCCCTGCAGAGAAAT	240	
Oy	989	GGACACAGTTCCTCAAGAGCGGGCGGTGGCAGTCTTGCCCCGCAACTGGCAGAGTCACTCA	1048	
Db	241	GGACACAGTTCCTCAAGAGCGGGCGGTGGCAGTCTTGCCCCGCAACTGGCAGAGTCACTCA	300	
Oy	1049	ACCAAGCTGACAGGCGATGACACACCTGACAGACACCTCTGGCAGACACCACTCTCTTG	1108	
Db	301	ACCAAGCTGACAGGCGATGACACACCTGACAGACACCTCTGGCAGACACCACTCTCTTG	360	
Oy	1109	GGGTTTTTCAAGCACAAGTGGGGGTACATATATCTGTGGGCACTATTTGAGTACCAAGTTGG	1168	
Db	361	GGGTTTTTCAAGCACAAGTGGGGGTACATATATCTGTGGGCACTATTTGAGTACCAAGTTGG	420	
Oy	1169	AAGAGATCCAGCAGGGGTGTTTGAAGGGCCCTTAAGAGAGTACCATGAGAGAGCCAGAAAT	1228	
Db	421	AAGAGATCCAGCAGGGGTGTTTGAAGGGCCCTTAAGAGAGTACCATGAGAGAGCCAGAAAT	480	
Oy	1229	GGGACCGCTACACTGACCTGTATACCAAGCCCTGGCCTGGCTGCTGCAATTAACAACTGGC	1288	

Db 481 GGGACGGCTACACTGACCCCTGTAACCAAGCCCTGGCCCTGCTGCATTAAACACTGGC 540

QY 1289 ATGGCGCCGACGAGCTACACAGCTCCCTG-GAGCTACCCGACAAATCCTCAACTTGGC 1347
|||||
Db 541 ATGGGGGCGACGGCTACACAGCTCCCTGTAAGCTACCCGACAAATCCTCAACTTGGC 600

QY 1348 AAGACGACCCGCTGATGAGAGAGAGAGTGGGCTCTGCTGAGAGCCGCCCTGCTGCTG 1407
|||||
Db 601 AAGACGACCCGCTGATGAGAGAGAGTGGGCTCTGCTGAGAGCCGCCCTGCTGCTG 658

QY 1408 AAGAGGGGACCA 1420
|||||
Db 659 AAGAAATGGCCCA 671

RESULT 7
BU518484 937 bp mRNA linear EST 12-SEP-2002
LOCUS AGENCOURT_10163679 NIH_MGC_134 Mus musculus cDNA clone
DEFINITION IMAGE:6516221 5', mRNA sequence.
ACCESSION BU518484
VERSION BU518484.1 GI:22826010
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: Dr. David Rowe
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Plate: L1AM14094 row: 9 column: 06
High quality sequence start: 19
High quality sequence stop: 671.
Location/Qualifiers
1. 937
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6516221"
/tissue_type="undifferentiated limb"
/lab_host="DH10B (phage-resistant)"
/clone_lid="NIH_MGC_134"
/note="Vector: pCMV-Sport6.1.ccd; Site.1: EcoRV; Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.7 kb. Constructed by Resgen, Invitrogen Corp. Note: this is a NIH_MGC library."

BASE COUNT 192 a 254 c 286 g 203 t 2 others

ORIGIN

Query Match 27.2% Score 621.2; DB 13; Length 937;
Best Local Similarity 85.4% Pzed. No. 9.9e-133;
Matches 703; Conservative 0; Mismatches 119; Indels 1; Gaps 1;

QY 890 GGGGAGTGGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGCA 949
|||||
Db 27 GGGGAGTGGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGCA 86

QY 950 AGGCGATATGGGGGCGGACGAGACCTGAGAGAGTGGACCACTTCTGTAAGGCGC 1009
|||||
Db 87 AGGTGATGATGGGGGCGGACGAGACCTGAGAGAGTGGACCACTTCTGTAAGGCTC 146

QY 1010 GCGTGGATGCTGCTGCCCCGAACTGGAGCTCTACTTCAACACAGCTGAGGCGATGCACA 1069
|||||

Db 147 GGTGGTGTCTAGACCCCTGACTGAGAGGTCTACTTCAACACAGCTGAAGCGGTGCACA 206

QY 1070 CCCTGACGAGACCTCTCTGGACAAACACACACTTCTTTGGGGTTTTCAGACAGTGGG 1129
|||||
Db 207 CCCTGAGGGGGGCTCTCTGGACAAACACACACTTCTTTGGGGTTTTCAGACAGTGGG 266

QY 1130 GTGACATGTAACCTGTGGCCATCTGTGATGACCAAGTGGAGAGATCCAGCGGTGTTG 1189
|||||
Db 267 GCGATATGAGACCTGTCTGAGTTGTGATGACCAAGTGGAGAGATCCAGCAAGTGTGTTG 326

QY 1190 AGGCCCCCTAATAGGATACCATGAGAGACCCGAGAGTGGAGCCGTACATGACCTG 1249
|||||
Db 327 AGGTCTCTTACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 386

QY 1250 TACCAGACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1309
|||||
Db 387 TACCAGACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 446

QY 1310 GCTCCCTGAGACTACCCGACAAATCTCAACTTGTGAGAGAGACCCGCTGATGAGG 1369
|||||
Db 447 GTTCCCTGAGAACTCCGAGACAAACCTCAACTTGTGAGAGAGACCCGCTGATGAGG 506

QY 1370 AGCAGGTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1429
|||||
Db 507 ACCAGGTGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 566

QY 1430 ACCGTGTGGGCGAGCGGTTACAGACTGTGAGAGACCACTTACAGTGTGCTGCTGCTGCTG 1489
|||||
Db 567 ACCTGTGTGGGCGAGCGGTTACAGACTGTGAGAGACCACTTACAGTGTGCTGCTGCTGCTG 626

QY 1490 GCACAGGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1549
|||||
Db 627 GTACAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 686

QY 1550 AGGAGCTGACGCTTTTGACAGAGAGCCCATGAGAACCTGTGCTGCTGCTGCTGCTGCTGCTG 1609
|||||
Db 687 AGGAGCTGACGCTTTTGACAGAGAGCCCATGAGAACCTGTGCTGCTGCTGCTGCTGCTGCTG 746

QY 1610 AGCTGCTCTTGTGCGGCTCCGCTCTGAGCTGTGAGAGCTGAGAGCTGAGAGCTGAGAGCTG 1669
|||||
Db 747 AGTGTCTCTTGTGCGGCTCCGCTCTGAGCTGTGAGAGCTGAGAGCTGAGAGCTGAGAGCTG 806

QY 1670 AGTA-TGCTCTGCTGAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1711
|||||
Db 807 AGTACCGCTTCTGTGTAACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 849

RESULT 8
BM926974 1394 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT_6682051 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5767867
DEFINITION 5', mRNA sequence.
ACCESSION BM926974
VERSION BM926974.1 GI:19377353
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Plate: L1AM12826 row: m column: 20


```

Db      181 AGGCTCCAGCTGAGAAATTTGAATGTACCCAGAAAGGAAAGGACACACAGCAAT 240
OY      494 GCTTCAACTTCAATCCGCTTCTCTGAGCCCTCAATAGCTCCACCTGTACGTGTGCA 553
Db      241 GCTTCAACTTCAATCCGCTTCTCTGAGCCCTCAATAGCTCCACCTGTATGTGTGGCA 300
OY      554 CCTACGCTTCCAGCCCACTGACCTACCTCAACATGTCTCACCCTTGTGGAGCATG 613
Db      301 CCTATGCTTCCAGCCCACTGACCTACCTCAACATGTCTCACCCTTGTGGAGCATG 360
OY      614 GAGATTTTGAAGTGGGAAAGGCAAGTGCCTATGACCAGATGAGGGGCACTGTGCGC 673
Db      361 CAGAAATTTGAGATGGGAAAGGCAAGTGCCTATGACCAGATGAGGGGCACTGTGCGC 420
OY      674 TTTCTGTGATGTGAGCTGTACTTCTGAGGACACTCAACATCTCTGAGGACGAAACCA 733
Db      421 TCTTGTGAGCGGTGAGCTGTACTTCTGAGGACACTCAACATCTCTGAGGACGAAACCA 480
OY      734 TTTATCTGCTTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 793
Db      481 TTTATCTGCTTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
OY      794 TCAGACACCTCACTTGTAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAG 853
Db      541 TGATATGAAACCCCTTGTAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAG 600
OY      854 GGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 913
Db      601 GAGACAGTGAACAAGATCTACTTCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
OY      914 CCGACAGTGTGTGCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 973
Db      661 CCGACAGTGTGTGCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
OY      974 CCTGTGAGAGAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1022
Db      721 CGCTGCAGAGAA-TGACAGACGTTCTGTGAGGCTGTGTGTGTGCTG 768

RESULT 10
LOCUS   BE270240      632 bp      mRNA      linear      EST 13-JUL-2000
DEFINITION BE270240 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:354355 5',
            mRNA sequence.
ACCESSION BE270240
VERSION   BE270240.1 GI:9143872
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE     Unpublished
JOURNAL   Contact: Robert Strausberg, Ph.D.
          Email: cgaaps-remail.nih.gov
COMMENT   Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
          CDNA Library Preparation: Ling Hong/Rubin Laboratory
          CDNA Library Arayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at: limgc.lnl.gov
          Plate: LINC237 row: f column: 08
          High quality sequence stop: 631.
          Location/Qualifiers
FEATURES
         source
         1..632
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="IMAGE:354355"
         /tissue_type="Burkitt lymphoma"
         /lab_host="DH10B (phage-resistant)"

```

```

BASE COUNT 131 a 185 c 192 g 124 t
ORIGIN
/c1one_11b="NIH_MGC_8"
/Note="Organ: Lymph; Vector: pOTB1; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCAAGAG(G). Size-selected >500bp for average
Insert size 1.8kb. Library constructed by Ling Hong in
the Laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
Query Match 26.1%; Score 597; DB 10; Length 632;
Best Local Similarity 99.5%; Pred. No. 3.4e-127;
Matches 630; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
OY      750 GGGGCCCCACCACTCCATGAAAGACAGATACCTGGCTTTGGCTCAACGAACTCACTT 809
Db      1 GGGGCCCCACCACTCCATGAAAGACAGATACCTGGCTTTGGCTCAACGAACTCACTT 60
OY      810 TGTAGGCTGTGCTATGATGAGTGTGGGAGCTTCAACGGGGAGAGAGAGT 869
Db      61 TGTAGGCTGTGCTATGATGAGTGTGGGAGCTTCAACGGGGAGAGAGT 120
OY      870 CTACTTCTTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 929
Db      121 CTACTTCTTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
OY      930 TCGTGTGGCCCTGTCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 989
Db      181 TCGTGTGGCCCTGTCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
OY      990 GACACAGTTCCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1049
Db      241 GACACAGTTCCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
OY      1050 CCAGCTCAGGCGATGACACACCTCTGAGAGACACCTCTGAGCAACACACCTCTTGTGG 1109
Db      301 CCAGCTCAGGCGATGACACACCTCTGAGAGACACCTCTGAGCAACACACCTCTTGTGG 360
OY      1110 GGTTTTCAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1169
Db      361 GGTTTTCAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
OY      1170 AGAGATCCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1229
Db      421 AGAGATCCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
OY      1230 -GGACCGCTACACTGACCTGTATCCAGCCCTGTGAGGCTGTGCTGATTAACTAGTGC 1288
Db      481 TGAGACCGCTACACTGACCTGTATCCAGCCCTGTGAGGCTGTGCTGATTAACTAGTGC 540
OY      1289 ATGCGGCGCAGCGCTACACCACTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1348
Db      541 ATGCGGCGCAGCGG-TACACCACTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 599
OY      1349 AGAAGCAACCGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1381
Db      600 AGAAGCA-CCGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 631

RESULT 11
LOCUS   B0068313      1002 bp      mRNA      linear      EST 02-APR-2002
DEFINITION B0068313 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5770567
            5', mRNA sequence.
ACCESSION B0068313
VERSION   B0068313.1 GI:19897359
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```


Db	445	GGGGTTCTCCAGACCGGCACTCCAGAAATTCCTGACACTGACGCTGAAGGAGGCCACTGG	504
OY	354	GCTTCTGACGTGGGCGCCGAGAGGCCCTGTTTGCCTTAGCAATGGAGGCCCTGGAGCT	413
Db	505	GCTTCTGACCTGGGCGCCGCCARAGGCCCTTTTGGCTTCAGCATGGAGGCCCTGGAGCT	564
OY	414	GCAAGGAGCGCATCTCTGGGAGGCCCGCGTGGACAAAGACATGATGATCCAGAAAG	473
Db	565	GCAAGGAGCGCATCTCTGGGAGGCCCGCGTGGACAAAGACATGATGATCCAGAAAG	624
OY	474	GAAAGACAACAGAGCCGAGTGCTTCAACTTCATCCGCTCCCTGCAGCCCTACAAATGCCTC	533
Db	625	GAAAGACAACAGAGCCGAGTGCTTCAACTTCATCCGCTCCCTGCAGCCCTACAAATGCCTC	684
OY	534	CCACCTGTAGCTGTGTGGCACCTACGCCCTTCAGACCCCAAGTGCACCTACGTCAACATGCT	593
Db	685	CCACCTGTAGCTGTGTGGCACCTACGCCCTTCAGACCCCAAGTGCACCTACGTCAACATGCT	744
OY	594	CACCTTCACTTTGGAGCATGGAGATTTGAAGATGGGAAGGGCAAGTCTCCCTATGACCC	653
Db	745	CACCTTCACTTTGGAGCATGGAGATTTGAAGATGGGAAGGGCAAGTCTCCCTATGACCC	804
OY	654	AGCTAAGGGCCATGCTGGCCTCTTGTGNGAATGATGAGCTGTACACCGGCC	702
Db	805	AGCTAAGGGCCATGCTGGCCTCTTGTGNGAATGATGAGCTGTACTCTGGGCC	853

RESULT 13	
BM962964	
LOCUS	BM962964 745 bp mRNA linear EST 18-MAR-2002
DEFINITION	UI-M-ECO-b.w1.c-05-0-ri r1 NIH.BMAP_ECO Mus musculus cDNA clone IMAGE:569876 5', mRNA sequence.
DESCRIPTION	bm962964

ACCESSION	BM962504	GI:19546384
VERSION	BM962504.1	
KEYWORDS	EST	
SOURCE	Mus musculus	(house mouse)
ORGANISM	Mus musculus	

REFERENCE	1 (bases 1 to 745)
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished
COMMENT	Contact: Robert Strausberg, Ph.D.

Email: c9agab3-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MCC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pyx-5.

FEATURES	Location/Qualifiers
source	1. .745

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone IMAGE:5698876"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_id="NH.BMAP.ECO"
/note="Organ: brain; Vector: pYX-Asc; Site.1: EcoR I;
Site.2: Not I; The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,

```

cloned with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is CAGCAGCAGC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

Query Match	24.9%	Score 569.2;	DB 12;	Length 745;
Best Local Similarity	85.6%	Pred. No. 9,46-121;		
Matches 631; Conservative	0;	Mismatches 106;	Indels 0;	Gaps 0;

O7 425 TCCTCCTGGGAGGCCCCCGTGGAGAAGAACTACTGTATCCAGAAAAGGGAACAACC 484
 ||| || | | | | | | | | | | | |
D6 7 TCTTTGGAGAGCTCCAGCTNGAGNAGAAATTGTGAATGTACCCAGAAAAGGGAAGCAACC 66

[illegible]

OY 545 TCCTGGCACCCTACGCCCTCCAGGCCAATGCAACTTACGTTCACAATGCCTACCTTTCACTT 604
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 TCTCGGCGACCTATGCTTTCAGAGCCAAATGCACTTAATCACATGCTCAAGTTGAACCT 186

QY	605	TGAGACATGGACAGTTTGAGAGATGGGAAGGGCAAGTGTCCCTATGACCCAGCTAAGGGCC	664
Db	187	TGGACCGTGCAGATTTGAGAGATGGGAAGGTTAATGCCCTATGACCCAGCTAAGGGTC	246

Oy		665	ATGCTGGCCTTCTGTGGATGAGTGTA	CTACTCGGCACACTGACAATTCTGGGCA	724
Db		247	ACACCGAGTCCTGTGGACGGTAGGTGT	ACTAGCCACACTGAATAACTCTTGGGCA	306

Dy 725 CGGAACCCATTATCCTGCGTAACATGG66CCCCACCACCTCCATGAGACAGACTCTGG 784
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 307 CAGAGCCGGTTATCCTTCGTATACATGGGAGCCCACCACTGCATTAAGACAGAGTACTGG 365

Oy 785 CTTTTGGCTACGAACTTCATTGTAGGCTCTGCCTATGTACTGAGAGTGTGGACA 844
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 367 CTTTITGGCTCAAGAACCACCTTGTAAGGCTCTGCCTTTGTCCCGAGAGTGTGGAAA 426

845 GCTTCACGGGGGACGACGACAGTCTACTCTCTTCACGGGAGCGGGCACTGGAGTCCG 904
 427 GCTTCACGGGGGAGAGATGACAAAGATCTACTCTCTCTTCACGACGCGGGCACTGGAGATG 486
 DB

905 ACTGCTATGCCGAGCAGCGTGTGGCTGTGTCGTCGACGAGCCCATATGCGGG 964
oy | |||||
db 487 ACTGCTATCCGACAGGTGTGGCTGTGTCGTCGACGAGCCCATATGCGGG 546
| |||||

[illegible][illegible]

1085 CCTGGCACACACACCTCTTTTGGGGTTTTCAAGCACAGTGGGTGACATGTACTGT 1144

CGGCCATCTGTGAGTAC 1161
|||||
737 GATCCTATTTCCTCATCTAC 743

DD 12/ C:\CSC011\BIOGRAPHIC /43
RESULT 14
0HE10004
0HE10004

LOCUS	BU612334	763 bp	mRNA	Linear	EST 20-FEB-2003
DEFINITION	UT-M-EMO-cax-k-14-0-UI.r1 NIH.BMAP-EMO Mus musculus cDNA clone UI-M-EMO-cax-k-14-0-UI 5', mRNA sequence.				

cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 130 a 176 c 175 g 116 t
ORIGIN

Query Match 23.5%; Score 536.4; DB 10; Length 597;
Best Local Similarity 98.3%; Pred. No. 3.3e-113;
Matches 584; Conservative 0; Mismatches 6; Indels 4; Gaps 4;

```
QY 1034 GGCAGCTCTACTTCAACCAAGCTGCAAGGATGACACACCTCGACAGACACTCTCTGGCACA 1093
DB 2 GGCAGCTCTACTTCAACCAAGCTGCAAGGATGACACACCTCGACAGACACTCTCTGGCACA 61
QY 1094 ACACACACTTCTTTGGGGTTTTCAGCAGACAGTGGGGTGACATGTACTGTGGCCATCT 1153
DB 62 ACACACACTTC-TTGGGGTTTATCAGCAGAGTGGGGTGACATGTACTGTGGCCATCT 120
QY 1154 GTGAGTACCACTTGGAAAGATCCAGGGGGTGTGAGGGGCCCTATAAGGAGTACCATG 1213
DB 121 GTGAGTACCAAG-TGGAAAGATCCAGGGGGTGTGAGGGGCCCTATAAGGAGTACCATG 179
QY 1214 AGGAAGCCCAAGATGGGACCGCTACACTGACCCCTGTACCCAGCCCTCGGCTGGCTCGT 1273
DB 180 AGGAAGCCCAAGATGGGACCGCTACACTGACCCCTGTACCCAGCCCTCGGCTGGCTCGT 239
QY 1274 GCATTACACTGGCATGGGCGCCACGGCTACACCAAGCTCCCTGGAGCTTACCCGACACA 1333
DB 240 GCATTACACTGGCATGGGCGCCACGGCTACACCAAGCTCCCTGGAGCTTACCCGACACA 299
QY 1334 TCCTCAACTTGTCAAGAAAGACCGCGCTGATGAGAGACAGTGGGGCTCGGTGGAGCC 1393
DB 300 TCCTCAACTTGTCAAGAAAGACCGCGCTGATGAGAGACAGTGGGGCTCGGTGGAGCC 359
QY 1394 GCCCCTGCTGCTGTAAGAAAGGACCAACTTCACCCACCTGCTGGGCGACCGGTTACAG 1453
DB 360 TGCCCTGCTGCTGTAAGAAAGGACCAACTTCACCCACCTGCTGGGCGACCGGTTACAG 419
QY 1454 GACTTGATGAGACCACTATACAGTGTGTTGATGAGCAGAGAGACGGCTGG-CTGCTC 1512
DB 420 GACTTGATGAGACCACTATACAGTGTGTTGATGAGCAGAGAGACGGCTGGACTGCTC 479
QY 1513 AAGGCTGTGAGCCTGGGGCCTGGGGTCACTGATTGAGAGCTGCAGCTG-TTGGACCA 1571
DB 480 AAGGCTGTGAGCCTGGGGCCTGGGGTCACTGATTGAGAGCTGCAGCTGTTGGACCA 539
QY 1572 GGAGCCCATGAGAAAGCCTGTGTATCTCAGAGCAAGAGCTGCTTTGGCGG 1625
DB 540 GGAGCCCATGAGAAAGCCTGTGTATCTCAGAGCAAGAGCTGCTTTGGCGG 593
```

Search completed: August 19, 2003, 15:52:41
Job time : 3213 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 12:26:06 ; Search time 437 Seconds

(without alignments)
14108.733 Million cell updates/sec

Title: us-10-002-050-13

Perfect score: 2284 1 cggcccttcacactctg.....tctgtgtgtgtgtgtgtgt 2284

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 segs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2280	99.8	2284	22	AAC84888 Human SEC7 nucleic
2	1887.4	82.6	2155	22	AAC84887 Human SEC6 nucleic
3	1879.4	82.3	2156	22	AAC84892 Human SEC11 nucleic
4	1761	77.1	3293	24	AAD08048 Human extracellular
5	1738	76.1	3868	24	AAD28947 Human MOL5a cDNA.
6	1722	75.4	3112	24	AAD28949 Human MOL5c cDNA.
7	1699	74.4	2558	24	AAD28948 Human MOL5b cDNA.
8	1617.6	70.8	3776	24	ABN63963 Human gene sequence

9	1394	61.0	8095	24	ABX92031 Lung specific nucl
10	1394	59.2	8144	25	ABT13397 Breast specific re
11	1351	59.2	3503	22	AAE29461 Murine M-Sema-F CD
12	1193.6	52.3	2811	22	AAE29451 Human TRNGO 276 CD
13	1094.2	47.9	2813	24	ABN59626 Novel human coding
14	928.4	40.6	1024	24	ABN85379 Partial human NOV2
15	841.4	36.8	967	22	AAE30193 Clone 14998905.0.6
16	463	20.3	729	22	AAE29452 Human TRNGO 276 CD
17	423	18.5	2132	24	ABL50334 Human cancer cell
18	411	18.0	3781	21	AAA37092 Human PRO1480 (UNC
19	411	18.0	3781	22	AAE54351 Human DNA encoding
20	411	18.0	3781	22	AAE54381 Primer #75 used in
21	411	18.0	3781	25	ACAS7909 Human PRO1480 cDNA
22	411	18.0	3781	25	ABX98379 Human cDNA encodin
23	411	18.0	3781	25	ABX98881 Novel human secret
24	411	18.0	3781	25	ACA05926 Human secreted/tra
25	411	18.0	3781	25	ABX87920 Human PRO polynuc
26	411	18.0	3781	25	ABX87924 Human PRO polynuc
27	411	18.0	3781	25	ABX75767 Human cDNA encodin
28	411	18.0	3781	25	ABX75767 Human PRO polynuc
29	411	18.0	3781	25	ABX16812 Human cDNA encodin
30	409.4	17.9	3556	22	AAD08283 Human secreted pro
31	409.4	17.9	3766	24	ABX77938 Hypoxia-induced pr
32	409.4	17.9	3766	24	ABN83984 Human gene sequenc
33	409.4	17.9	3766	25	ACCS1059 Human bladder canc
34	409.4	17.9	3766	25	ABX76369 Lung cancer-associ
35	409	17.9	2703	21	AAA47437 Sequence encoding
36	406.8	17.8	2390	24	ABR11109 DNA encoding human
37	401	17.6	4157	18	AAE50665 Human CD100 antige
38	401	17.6	4157	22	AAE77463 Human CD100 DNA se
39	401	17.6	4157	22	AAE77463 Human CD100 nucle
40	398.2	17.4	2769	19	AAV31121 Human semaphorin e
41	398.2	17.4	2769	22	AAE77462 Murine CD100 DNA s
42	398.2	17.4	2769	22	AAE77462 Mouse CD100 nucle
43	398.2	17.4	4391	18	AAE50665 Mouse CD100 antige
44	370.4	16.2	2517	24	ABN4076 Human semaphorin G
45	370.4	16.2	2598	24	ABA04078 Human semaphorin G

ALIGNMENTS

RESULT 1	
AAC84888	
ID	AAC84888 standard; cDNA; 2284 BP.
XX	
AC	AAC84888;
XX	
DT	20-APR-2001 (first entry)
XX	
DE	Human SEC7 nucleic acid sequence (clone ID 20422974.2).
XX	
SE	SEC6; cytosolic; gynecological; gene therapy; screening assay; human;
KW	chromosomal mapping; forensic biology; cell proliferation; cancer;
KW	cell differentiation; immune associated disorder; gestational disease;
KW	SEC6; ss.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	/*tag="a
FT	/product="SEC"
XX	
PN	MO200078802-A2.
XX	
PD	28-DEC-2000.
XX	
PF	23-JUN-2000; 2000MO-US17328.
XX	
PR	23-JUN-1999; 99US-0140584.
PR	20-JUL-1999; 99US-0144720.
PR	16-SEP-1999; 99US-0154520.

22-JUN-2000; 2000US-0604286.
 XX (CURA-) CURAGEN CORP.
 XX Shinkens RA, Fernandes E, Vernet C, Yang M, Boldog FL;
 PI Herrmann JI;
 XX WPI; 2001-071385/08.
 DR P-PSDB; AAB48374.
 XX Polynucleotides encoding SECX proteins useful for treating disease
 PT characterized by an aberrant level of cell proliferation and/or
 PI differentiation like cancer or immune associated disorders -
 XX
 PS Claim 3; Fig 7; 132p; English.
 CC The invention relates to human SECX polypeptides and polynucleotides
 CC encoding them. The SECX polypeptides can be expressed by standard
 CC recombinant methodology. The SECX polypeptides are useful for treating
 CC or preventing a SECX-associated disorder. The invention is useful in
 CC screening assays; detection assays (e.g. chromosomal mapping, cell and
 CC tissue typing, forensic biology); predictive medicine (diagnostic assays,
 CC prognostic assays, monitoring clinical trials, and pharmacogenomics); and
 CC methods of treatment (e.g. therapeutic and prophylactic), especially
 CC disorders characterized by aberrant cell proliferation and/or
 CC differentiation like cancer or immune associated disorders or gestational
 CC disease. The present sequence represents a SEC7 nucleic acid sequence.
 CC
 SQ Sequence 2284 BP; 443 A; 683 C; 689 G; 465 T; 4 other;
 Query Match 99.8%; Score 2280; DB 22; Length 2284;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 601 ACTTTGAGCAGAGAGATTGAAAGTGGAAAGGCAAGTGTCCATATGACCCAGCTAAG 660
 DB 601 ACTTTGAGCAGAGAGATTGAAAGTGGAAAGGCAAGTGTCCATATGACCCAGCTAAG 660
 QY 661 GGCATGCTGGCCTTCTTGTGATGAGTGTACTGGCCCACTCAACAATCTTCG 720
 DB 661 GGCATGCTGGCCTTCTTGTGATGAGTGTACTGGCCCACTCAACAATCTTCG 720
 QY 721 GGCAGGAAACCATTAATCTGGCTAAATGAGGGGCCCCACACTCCATGAAGAGATAC 780
 DB 721 GGCAGGAAACCATTAATCTGGCTAAATGAGGGGCCCCACACTCCATGAAGAGATAC 780
 QY 781 CTGGCTTTTGGCTCAACGAAACCTCACTTGTAGGCTGTGCTATGATCTAGAGATG 840
 DB 781 CTGGCTTTTGGCTCAACGAAACCTCACTTGTAGGCTGTGCTATGATCTAGAGATG 840
 QY 841 GGCAGCTTACGGGGGACGACACAAAGTCTACTTCTTTCAGGAGCGGGCATGTAG 900
 DB 841 GGCAGCTTACGGGGGACGACACAAAGTCTACTTCTTTCAGGAGCGGGCATGTAG 900
 QY 901 TCCGACTGTATGCGAGGAGGAGGTGGCTGATGAGCCCGTGTGCAAGGGCGATATG 960
 DB 901 TCCGACTGTATGCGAGGAGGAGGTGGCTGATGAGCCCGTGTGCAAGGGCGATATG 960
 QY 961 GGGGGGACGAGACCTGTGAGAGAGTGGACACGTTCTTAAGGCGGGCTGGCATC 1020
 DB 961 GGGGGGACGAGACCTGTGAGAGAGTGGACACGTTCTTAAGGCGGGCTGGCATC 1020
 QY 1021 TCTGCCCGCAATGGAGAGTCTACTTCAACAGCTGACAGGCACTGACACCTGACAG 1080
 DB 1021 TCTGCCCGCAATGGAGAGTCTACTTCAACAGCTGACAGGCACTGACACCTGACAG 1080
 QY 1081 ACCCTGTGACAAACACACCTTCTTGGGGTTTTCAAGCAAGTGGGGTGCATGTAC 1140
 DB 1081 ACCCTGTGACAAACACACCTTCTTGGGGTTTTCAAGCAAGTGGGGTGCATGTAC 1140
 QY 1141 CTGTGGCCATGTGTAGTACAGATTGGAAGATTCACGCGGTGTTTGAAGGCCCTAT 1200
 DB 1141 CTGTGGCCATGTGTAGTACAGATTGGAAGATTCACGCGGTGTTTGAAGGCCCTAT 1200
 QY 1201 AAGGATGATCATGAGGAAACCCAGAAAGTGGACCGTATACAGACCTTACCCAGCCT 1260
 DB 1201 AAGGATGATCATGAGGAAACCCAGAAAGTGGACCGTATACAGACCTTACCCAGCCT 1260
 QY 1261 CGGCTGTGCTGTGATTAACAATGTGCAATCGGCGCACGCGTACACAGTCCCTGAG 1320
 DB 1261 CGGCTGTGCTGTGATTAACAATGTGCAATCGGCGCACGCGTACACAGTCCCTGAG 1320
 QY 1321 CTACCCGACAAATCTCAACTTCGTCAAGAAAGACCCGCTGATGAGAGAGAGTGGG 1380
 DB 1321 CTACCCGACAAATCTCAACTTCGTCAAGAAAGACCCGCTGATGAGAGAGAGTGGG 1380
 QY 1381 CTTGGTGAAGCGCCCGCTGCTGCGTGAAGAAAGGACCAACTTCAACCACTGTGGGC 1440
 DB 1381 CTTGGTGAAGCGCCCGCTGCTGCGTGAAGAAAGGACCAACTTCAACCACTGTGGGC 1440
 QY 1441 GACCGGGTTACAGACTTGTATGAGACCACTTATACAGTCTGTTCATTGGCAGAGAG 1500
 DB 1441 GACCGGGTTACAGACTTGTATGAGACCACTTATACAGTCTGTTCATTGGCAGAGAG 1500
 QY 1501 GGTGGCTGCTGAAGGCTGTAGCCTGGGGCCCTGGGTTCACTGATGAGAGCTGAG 1560
 DB 1501 GGTGGCTGCTGAAGGCTGTAGCCTGGGGCCCTGGGTTCACTGATGAGAGCTGAG 1560
 QY 1561 CTGTTTGAACAGAGACCCATGAAAGCCGTGGTGTATCTCAAGAGCAAGAGTCTTT 1620
 DB 1561 CTGTTTGAACAGAGACCCATGAAAGCCGTGGTGTATCTCAAGAGCAAGAGTCTTT 1620
 QY 1621 GCGGCTCCCGCTCAGCTGAGTGGCGGCTGGCGGAGCTGCAATGAAGATGAGCTTC 1680
 DB 1621 GCGGCTCCCGCTCAGCTGAGTGGCGGCTGGCGGAGCTGCAATGAAGATGAGCTTC 1680

QY		1681	TGTCACAGCTGTGCTCGCCCGGAGCCCTATTGGCCGTGAGACGTCAACACAGCCGC	1740
Db		1681	TGTGCACAGCTGTGCTCGCCCGGAGCCCTATTGGCCGTGAGACGTCAACACAGCCGC	1740
QY		1741	TGTGTGCGCCGTGCGTGGCCACTTTGGATCTTTACTGATCCAGCATGTGATGACCTCGAC	1800
Db		1741	TGTGTGCGCCGTGCGTGGCCACTTTGGATCTTTACTGATCCAGCATGTGATGACCTCGAC	1800
QY		1801	ACTTCAGGACATTTCCACACTCCCGGAGAGTAAGATACAGTACAGGCCACATNCCAAAAC	1860
Db		1801	ACTTCAGGACATTTCCACACTCCCGGAGAGTAAGATACAGTACAGGCCACATNCCAAAAC	1860
QY		1861	ATTCAGGCTGTGTGCGGGCAGACACTGTGCTGCCCTGCCACACCTCTCCTCCAACTTGGCC	1920
Db		1861	ATTCAGGCTGTGTGCGGGCAGACACTGTGCTGCCCTGCCACACCTCTCCTCCAACTTGGCC	1920
QY		1921	CTGCCCCAGCTCCAAACCCCGAGAGTATCATGATAGAGGGAAACCCCAACGCGTGGCGG	1980
Db		1921	CTGCCCCAGCTCCAAACCCCGAGAGTATCATGATAGAGGGAAACCCCAACGCGTGGCGG	1980
QY		1981	ANACGCTGGAGGTGATGACTCTACTTTGGACAGGACACAGCTATCTCAGAGGACATGGC	2040
Db		1981	ANACGCTGGAGGTGATGACTCTACTTTGGACAGGACACAGCTATCTCAGAGGACATGGC	2040
QY		2041	ACGGGCACCTGCTGTCTGTGGACAGATACTGCCCCAGCACCCACCGCGGCATYGAGACT	2100
Db		2041	ACGGGCACCTGCTGTCTGTGGACAGATACTGCCCCAGCACCCACCGCGGCATYGAGACT	2100
QY		2101	GCTGTGCGACAGAGGGGACATGCACCTTGTTGTTGTCACAGGGACACAGCTCGGAGAAAG	2160
Db		2101	GCTGTGCGACAGAGGGGACATGCACCTTGTTGTTGTTGTCACAGGGACACAGCTCGGAGAAAG	2160
QY		2161	CATCTTCCTCTCTCTGTGATTCACAGACAGCGCGGACCCAGCCGCGCAAAATTTTCAAG	2220
Db		2161	CATCTTCCTCTCTCTGTGATTCACAGACAGCGCGGACCCAGCCGCGCAAAATTTTCAAG	2220
QY		2221	GCAGAAAGTTNAAATGTGTGTTTGTGATTTGACATGTGTTGTGTGTGTATGTG	2280
Db		2221	GCAGAAAGTTNAAATGTGTGTTTGTGATTTGACATGTGTTGTGTGTGTATGTG	2280
QY		2281	TGTG	2284
Db		2281	TGTG	2284
RESULT 2				
AAAC84887				
ID	AAAC84887	standard; cdna; 2155 BP.		
XX				
AC	AAAC84887;			
XX				
DT	20-Apr-2001	(first entry)		
XX				
DE		Human SEC6 nucleic acid sequence (clone ID 20422974.0.132).		
XX				
KM	SEC6; cytosolic; gynecological; gene therapy; screening assay; human;			
KM	chromosomal mapping; forensic biology; cell proliferation; cancer;			
KM	cell differentiation; Immune associated disorder; gestational disease;			
KM	SEC6; ss.			
XX				
OS	Homo sapiens.			
XX				
EH	Key	Location/Qualifiers		
FT	CDS	166..1937		
FT		/*tag= a		
FT		/product= "SEC6"		
XX				
PN	MO200078802-A2.			
XX				
XX	28-DEC-2000.			
XX				
PF	23-JUN-2000; 2000MO-DS17328.			
XX				

PR	23-JUN-1999;	99US-0140584.
PR	20-JUN-1999;	99US-0144722.
PR	16-SEP-1999;	99US-0154520.
PR	22-JUN-2000;	2000US-0604286.
XX		
PA	(CURA-) CURAGEN CORP.	
XX		
PI	Shimkets RA, Fernandes E, Vernet C, Yang M, Boldog FL;	
PI	Herrmann JL;	
DR	WPI: 2001-071385/08.	
XX	P-PSDB; AAB48373.	
XX		
PT	Polynucleotides encoding SECX proteins useful for treating disease	
PT	characterized by an aberrant level of cell proliferation and/or	
PT	differentiation like cancer or immune associated disorders	
XX		
PS	Claim 3; Fig 6; 132pp; English.	
XX		
CC	The invention relates to human SECX polypeptides and polynucleotides	
CC	encoding them. The SECX polypeptides can be expressed by standard	
CC	recombinant methodology. The SECX polypeptides are useful for treating	
CC	or preventing a SECX-associated disorder. The invention is useful in	
CC	screening assays; detection assays (e.g. chromosomal mapping, cell and	
CC	tissue typing, forensic biology); predictive medicine (diagnostic assays, and	
CC	prognostic assays, monitoring clinical trials, and pharmacogenomics); and	
CC	methods of treatment (e.g. therapeutic and prophylactic), especially	
CC	disorders characterized by aberrant cell proliferation and/or	
CC	differentiation like cancer or immune associated disorders or gestational	
CC	disease. The present sequence represents a SEC6 nucleic acid sequence.	
XX		
SO	Sequence 2155 BP; 444 A; 614 C; 645 G; 452 T; 0 other;	
	Query Match 82.6%; Score 1887.4; DB 22; Length 2155;	
	Best Local Similarity 99.5%; Pred. No. 0;	
	Matches 1914; Conservative 0; Mismatches 6; Indels 3; Gaps 2	
OY	1	CGGCCCTTCTACACTCTCGCCCTGCTGATGNGAAGGGGTTGGGGTTCTGCAGGGCT 60
DB	1	CGGCCCTTCTACACTCTCGCCCTGCTGATGNGAAGGGGTTGGGGTTCTGCAGGGCT 60
OY	61	ATTGTCTGCGTGGGGAAGGGGACAGGCCGGGACCTCGCTCGACGCGCGCG 120
DB	61	ATTGTCTGCGTGGGGAAGGGGACAGGCCGGGACCTCGCTCGACGCGCGCG 120
OY	121	ACGACGAGAGACAGTGGGCTTAAGCTAGAGCGCGGGGGTGGGCCATGGCCCCACACTGG 180
DB	121	ACGACGAGAGACAGTGGGCTTAAGCTAGAGCGCGGGGGTGGGCCATGGCCCCACACTGG 180
OY	181	GCTGCTGGCTGCTGGAGCAAGAGGTGTGGGGCTGGGGCATTTGGGGCTGAGGTGTGTGG 240
DB	181	GCTGCTGGCTGCTGGAGCAAGAGGTGTGGGGCTGGGGCATTTGGGGCTGAGGTGTGTGG 240
OY	241	AACCTTGTGCGCGGTAAAGACAGTGTCTTCTGGGGAGCTGGCCACGGTATGAGCGCGTTT 300
DB	241	AACCTTGTGCGCGGTAAAGACAGTGTCTTCTGGGGAGCTGGCCACGGTATGAGCGCGTTT 300
OY	301	TCCGACAGACCGGACATCAGAGACTTCTGTACATGACGTGACGAGGACCCACTGGGCTTCTG 360
DB	301	TCCGACAGACCGGACATCAGAGACTTCTGTACATGACGTGACGAGGACCCACTGGGCTTCTG 360
OY	361	TACGNGGCGCGCCGAGAGGCCCTGTGTTGCTTCAAGATGAGAGGCCCTGAGAGCTGCAAGA 420
DB	361	TACGNGGCGCGCCGAGAGGCCCTGTGTTGCTTCAAGATGAGAGGCCCTGAGAGCTGCAAGA 420
OY	421	GCGATCTCTTGGAGAGCCCGCTGGAGAAAGAACTGATGATCCAGAAAGGAAGAAC 480
DB	421	GCGATCTCTTGGAGAGCCCGCTGGAGAAAGAACTGATGATCCAGAAAGGAAGAAC 480
OY	481	AACGACAGCAGAGTGTTCACATTAATATCGCTTCGACAGCCCTAATATCCCTCCACCTG 540
DB	481	AACGACAGCAGAGTGTTCACATTAATATCGCTTCGACAGCCCTAATATCCCTCCACCTG 540

```

OY 541 TACGCTGTGGACCTAGGCTTCCAGCCCAAGTGTACCTAGCTCAACATGCTCACCTTC 600
DB 541 TACGCTGTGGACCTAGGCTTCCAGCCCAAGTGTACCTAGCTCAACATGCTCACCTTC 600
OY 601 ACTTTGAGCATGAGAGATTGAGATGGAGAGGCAAGTGTCCCTATGACCCAGCTAAG 660
DB 601 ACTTTGAGCATGAGAGATTGAGATGGAGAGGCAAGTGTCCCTATGACCCAGCTAAG 660
OY 661 GGGCATTGCTGCGCTTCTGTGGATGGTGAAGTGTACTGGGCAACACTCAACTTCCTG 720
DB 661 GGGCATTGCTGCGCTTCTGTGGATGGTGAAGTGTACTGGGCAACACTCAACTTCCTG 720
OY 721 GGCAGGAAACCCATTATCCTGGGTAAACATGGGGGCCCCACACTCCATGAAAGCAAGTAC 780
DB 721 GGCAGGAAACCCATTATCCTGGGTAAACATGGGGGCCCCACACTCCATGAAAGCAAGTAC 780
OY 781 CTGGCTTTTGGCTCAAGCAACCTCACTTTTGAAGCTCTGCTATGATGACCTGAGAGTGG 840
DB 781 CTGGCTTTTGGCTCAAGCAACCTCACTTTTGAAGCTCTGCTATGATGACCTGAGAGTGG 840
OY 841 GGCACCTTCAGGGGGGAGCAGACAGCAAGGTCTACTTCTTTCAGGAGCGGGCAGTGGAG 900
DB 841 GGCACCTTCAGGGGGGAGCAGACAGCAAGGTCTACTTCTTTCAGGAGCGGGCAGTGGAG 900
OY 901 TCCGACTGTATGCGGAGAGGTGGTGGCTGCTGTGGGCCCCGTCTGCAAGGGCGCATATG 960
DB 901 TCCGACTGTATGCGGAGAGGTGGTGGCTGCTGTGGGCCCCGTCTGCAAGGGCGCATATG 960
OY 961 GGGGGCGCAGGAGCCCTGAGAGAGAACTGAGACACGTTCTGTAAGCGCGCGCTGGCATG 1020
DB 961 GGGGGCGCAGGAGCCCTGAGAGAGAACTGAGACACGTTCTGTAAGCGCGCGCTGGCATG 1020
OY 1021 TCTGCCCGCAACTGGCAGCTCTACTTCAACACAGCTGCAAGGCGATCACACCTGCAAGAC 1080
DB 1021 TCTGCCCGCAACTGGCAGCTCTACTTCAACACAGCTGCAAGGCGATCACACCTGCAAGAC 1080
OY 1081 ACCCTGGGCAACACCACTCTTTGGGGTTTTCAGGACAGTGGGGTGACATGTAC 1140
DB 1081 ACCCTGGGCAACACCACTCTTTGGGGTTTTCAGGACAGTGGGGTGACATGTAC 1140
OY 1141 CTGTGCGGCATCTGTGATACCAAGTGTGAAAGATTCACAGCGGGGTGTTGAGGGCCCTAT 1200
DB 1141 CTGTGCGGCATCTGTGATACCAAGTGTGAAAGATTCACAGCGGGGTGTTGAGGGCCCTAT 1200
OY 1201 AAGGATACCATGAGGAAAGCCAGAAAGTGGACCCCTACACTGACCTGTACCCAGCCCT 1260
DB 1201 AAGGATACCATGAGGAAAGCCAGAAAGTGGACCCCTACACTGACCTGTACCCAGCCCT 1260
OY 1261 CGGCTGGCTGTGATTAACAACTGGCATGCGGCGCACAGGCTACACGACCTCCCTGGAG 1320
DB 1261 CGGCTGGCTGTGATTAACAACTGGCATGCGGCGCACAGGCTACACGACCTCCCTGGAG 1320
OY 1321 CTACCCGCAACATCTCACTTCGTCAAGAAACACCCGCTGATGAGAGAGAGGTGGAG 1380
DB 1321 CTACCCGCAACATCTCACTTCGTCAAGAAACACCCGCTGATGAGAGAGAGGTGGAG 1380
OY 1381 CCTGGGTGAGCGCGCCCTGCTGTGAAGAGGACCAACTTCAACCACTGCTGTGGCC 1440
DB 1381 CCTGGGTGAGCGCGCCCTGCTGTGAAGAGGACCAACTTCAACCACTGCTGTGGCC 1440
OY 1441 GACCGGTTAGAGACTGATGAGACCACTATACAGTGTGTTTCAATGGGACAGAGAGAC 1500
DB 1441 GACCGGTTAGAGACTGATGAGACCACTATACAGTGTGTTTCAATGGGACAGAGAGAC 1500
OY 1501 GGGTGGCTCTCAAGGCTGTGAGCGCTGGGCGCTGGTCAACCTGATGAGAGAGCTGGAG 1560
DB 1501 GGGTGGCTCTCAAGGCTGTGAGCGCTGGGCGCTGGTCAACCTGATGAGAGAGCTGGAG 1560
OY 1561 CTGTTTGAACAGAGAGCCATGAGAAAGCTGTGCTATCTCAGAGCAAGAGCTGCTTTT 1620
DB 1561 CTGTTTGAACAGAGAGCCATGAGAAAGCTGTGCTATCTCAGAGCAAGAGCTGCTTTT 1620
OY 1621 GCCGGCTCCCGTCTCAGCTGGTGAAGTGGCCCGTGGCGCACTGATAAAGTATGCTCC 1680

```

```

DB 1621 GCCGGCTCCCGTCTCAGCTGGTGAAGCTGCGCGCGAGCTGCAATAAGTATGCTCC 1680
OY 1681 TGTGACAGCTGTGCTCGCCGCGGACCCCTATTGGGCGCTGGAGGCTCAACAGCCGCG 1740
DB 1681 TGTGACAGCTGTGCTCGCCGCGGACCCCTATTGGGCGCTGGAGGCTCAACAGCCGCG 1740
OY 1741 TGTGTGGCCGTGGGCGCACCTTTGATCTTACTGATACACATGATGATGACCTGAGC 1800
DB 1741 TGTGTGGCCGTGGGCGCACCTTTGATCTTACTGATACACATGATGATGACCTGAGC 1800
OY 1801 ACTTCAGGATTTGCAACCTCCGTGGCAGTAAAGATACATGACAGCCACTTCCAAAAC 1860
DB 1801 ACTTCAGGATTTGCAACCTCCGTGGCAGTAAAG - AAGTCAAGCCCACT - CCAAAAAC 1857
OY 1861 ATCAGGCTGTGGGCGCACAGACCGGTGTCGCCCTGCCACCTGCTCCCAACTTGGCC 1920
DB 1861 ATCAGGCTGTGGGCGCACAGACCGGTGTCGCCCTGCCACCTGCTCCCAACTTGGCC 1917
OY 1921 CTG 1923
DB 1918 CCG 1920

RESULT 3
AAC84892
ID AAC84892 standard; cDNA; 2156 BP.
XX AAC84892;
AC 20-APR-2001 (first entry)
XX
DT 20-APR-2001 (first entry)
XX
DE Human SEC11 nucleic acid sequence (clone ID 20422974.0.132-ext2).
XX
SEXC: cytostatic; gynecological; gene therapy; screening assay; human;
KW chromosomal mapping; forensic biology; cell proliferation; cancer;
KW cell differentiation; immune associated disorder; gestational disease;
XX SEC11; 58.
XX
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 166..2040
XX FT /'lag= a
XX FT /product= "SEC11"
XX
XX W0200078802-A2.
XX
XX 28-DEC-2000.
XX
XX 23-JUN-2000; 2000MO-US17328.
XX
XX 23-JUN-1999; 99US-0140584.
XX 20-JUL-1999; 99US-0144722.
XX 16-SEP-1999; 99US-0154520.
XX 22-JUN-2000; 2000US-0604286.
XX
XX (CUBA-) CUBAGEN CORP.
XX
XX Shinkels RA, Fernandes E, Vernet C, Yang M, Boldog FL,
XX Herrmann JI,
XX
XX WPI: 2001-071385/08.
XX P-PSDB: AAB48378.
XX
XX polynucleotides encoding SECX proteins useful for treating disease
XX characterized by an aberrant level of cell proliferation and/or
XX differentiation like cancer or immune associated disorders -
XX
XX Claim 3; Fig 11; 132pp; English.
XX
XX The invention relates to human SECX polypeptides and polynucleotides
XX encoding them. The SECX polypeptides can be expressed by standard

```

CC recombinant methodology. The SECX polypeptides are useful for treating
CC or preventing a SECX-associated disorder. The invention is useful in
CC screening assays (detection assays (e.g. chromosomal mapping, cell and
tissue typing, forensic biology); predictive medicine (diagnostic assays,
CC prognostic assays, monitoring clinical trials, and pharmacogenomics); and
CC methods of treatment (e.g. therapeutic and prophylactic), especially
CC disorders characterized by aberrant cell proliferation and/or
CC differentiation like cancer or immune associated disorders or gestational
CC disease. The present sequence represents a SEC11 nucleic acid sequence.

XX
SQ Sequence 2156 BP; 444 A; 617 C; 645 G; 450 T; 0 other;

Query Match 82.3%; Score 1879.4; DB 22; Length 2156;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 1909; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

QY 1 CGGCCCTTCCACACCTCCCTGCTGATGGAACGGGGTTGGGGTCTCAGAGGCT 60
DB 1 CGGCCCTTCCACACCTCCCTGCTGATGGAACGGGGTTGGGGTCTCAGAGGCT 60
QY 61 ATTGTCGCGCTGGGGAAGGGGACAGGCGGGAACCGGACCTCGCTCGGAGCGGCGC 120
DB 61 ATTGTCGCGCTGGGGAAGGGGACAGGCGGGAACCGGACCTCGCTCGGAGCGGCGC 120
QY 121 ACCAGAGGACAGCTGGCTGAAAGCTCAGACCGGGGCGTGGCCATGGCCCACTGG 180
DB 121 ACCAGAGGACAGCTGGCTGAAAGCTCAGACCGGGGCGTGGCCATGGCCCACTGG 180
QY 181 GCGTCGCGCTGGGAGGAGGAGGCTGGGGCTGGGGCATGGGGCTGAGGTGTGG 240
DB 181 GCGTCGCGCTGGGAGGAGGAGGCTGGGGCTGGGGCATGGGGCTGAGGTGTGG 240
QY 241 AACCTGTGCGGCTGAGAGAGAGTGTCTTGGGGAGCTGGCCAGGTAGTACGGCGTTC 300
DB 241 AACCTGTGCGGCTGAGAGAGAGTGTCTTGGGGAGCTGGCCAGGTAGTACGGCGTTC 300
QY 301 TCCCAAGCCGGCATCCAGAGCTTCCTGACACTGACGTCGAGGCCACATGGGCTTCTG 360
DB 301 TCCCAAGCCGGCATCCAGAGCTTCCTGACACTGACGTCGAGGCCACATGGGCTTCTG 360
QY 361 TAGCTGGGCGCCGAGAGGCGCTGTTGCTTCGATGAGGAGCGCTGAGCTGCAAGGA 420
DB 361 TAGCTGGGCGCCGAGAGGCGCTGTTGCTTCGATGAGGAGCGCTGAGCTGCAAGGA 420
QY 421 GCGATCTCTGGAGAGGCGCCCGTGGAGAGAAAGACTGATATCCAGAAAGGAGAAC 480
DB 421 GCGATCTCTGGAGAGGCGCCCGTGGAGAGAAAGACTGATATCCAGAAAGGAGAAC 480
QY 481 AACCAAGCAGAGTCTTCAACTTCATCGCTTCTGAGCCCTACATGGCTCCACCTG 540
DB 481 AACCAAGCAGAGTCTTCAACTTCATCGCTTCTGAGCCCTACATGGCTCCACCTG 540
QY 541 TAGCTCTGGGACCTTACGCTTCCAGCCCAAGTGCATACGTAACATGCTTACTTTC 600
DB 541 TAGCTCTGGGACCTTACGCTTCCAGCCCAAGTGCATACGTAACATGCTTACTTTC 600
QY 601 ACTTTGGAGCATGAGAGTGTGAAGATGGGAGGGAAGTGTCCCTATGACCCAGTAAG 660
DB 601 ACTTTGGAGCATGAGAGTGTGAAGATGGGAGGGAAGTGTCCCTATGACCCAGTAAG 660
QY 661 GGGCAGCTGCGCTTCTTGGATGGTGTGAGCTGACTCGGCCACATCAACTTCTG 720
DB 661 GGGCAGCTGCGCTTCTTGGATGGTGTGAGCTGACTCGGCCACATCAACTTCTG 720
QY 721 GGCAGGGAACCATATATCTGCGCTTAACATGGGCGCCCAACATCTCATGAAGAGTAC 780
DB 721 GGCAGGGAACCATATATCTGCGCTTAACATGGGCGCCCAACATCTCATGAAGAGTAC 780
QY 781 CTGGGCTTTTGGCTCAACGAACCTCACTTTGATGCTGTGCTATGTAAGTGTG 840
DB 781 CTGGGCTTTTGGCTCAACGAACCTCACTTTGATGCTGTGCTATGTAAGTGTG 840
QY 841 GGCAGCTTCAAGGGGAGCAGACAGTCTTCTTCAAGGAGCGGGCAGTGGAG 900

DB 841 GGCAGCTTCAAGGGGAGCAGACAGTCTTCTTCAAGGAGCGGGCAGTGGAG 900
QY 901 TCCGACTCATATGCGGAGAGTGTGGGCTGCTGGGCGGCTGCTGCAAGGGGATATG 960
DB 901 TCCGACTCATATGCGGAGAGTGTGGGCTGCTGGGCGGCTGCTGCAAGGGGATATG 960
QY 961 GGGGGCCAGCAGACCTTCAGAGAGAGTGCACACGTTCTGAAAGCGCGCTGGCATTC 1020
DB 961 GGGGGCCAGCAGACCTTCAGAGAGAGTGCACACGTTCTGAAAGCGCGCTGGCATTC 1020
QY 1021 TGTGCCCCGAATGCGAGCTCTACTTCAACAGTGCAGGCGATGACACCTCGACAGAC 1080
DB 1021 TGTGCCCCGAATGCGAGCTCTACTTCAACAGTGCAGGCGATGACACCTCGACAGAC 1080
QY 1081 ACCCTCCGAGCAACACACCTTCTTGGGGTTTCAACACAGTGGGGTGCATGTAC 1140
DB 1081 ACCCTCCGAGCAACACACCTTCTTGGGGTTTCAACACAGTGGGGTGCATGTAC 1140
QY 1141 CTGTGCGCATCTGTGAGTACCAAGTGAAGAGATCCAGCGGGTGTGAGAGGCCCTAT 1200
DB 1141 CTGTGCGCATCTGTGAGTACCAAGTGAAGAGATCCAGCGGGTGTGAGAGGCCCTAT 1200
QY 1201 AAGGATACCATATGAGAGAGCCCAAGAGTGGAGCCGTACACTGACCTGTACCGGCT 1260
DB 1201 AAGGATACCATATGAGAGAGCCCAAGAGTGGAGCCGTACACTGACCTGTACCGGCT 1260
QY 1261 CGGCTGCGCTGCTGATTAACAACCTGCGAGCGGCGACAGGCTACACACCTCCGTGAG 1320
DB 1261 CGGCTGCGCTGCTGATTAACAACCTGCGAGCGGCGACAGGCTACACACCTCCGTGAG 1320
QY 1321 CTACCCGACAAACATCTCAACTTCTGCAAGAGCAGCCGCTGATGAGAGAGAGTGGG 1380
DB 1321 CTACCCGACAAACATCTCAACTTCTGCAAGAGCAGCCGCTGATGAGAGAGAGTGGG 1380
QY 1381 CTTGCGGTGAGAGCGCCCTGCTGCTGTAAGAGAGGCAACCACTTACCCACTGCTG 1440
DB 1381 CTTGCGGTGAGAGCGCCCTGCTGCTGTAAGAGAGGCAACCACTTACCCACTGCTG 1440
QY 1441 GACCGGCTTACAGAGCTTGAAGAGCAGCCTATGAGTGTGCTTATGGACAGAGAG 1500
DB 1441 GACCGGCTTACAGAGCTTGAAGAGCAGCCTATGAGTGTGCTTATGGACAGAGAG 1500
QY 1501 GCGTGGCTGCTCAAGGCTGTGAGGCTGGGGCGTGGGTTGACCTGATGAGAGACTCAG 1560
DB 1501 GCGTGGCTGCTCAAGGCTGTGAGGCTGGGGCGTGGGTTGACCTGATGAGAGACTCAG 1560
QY 1561 CTGTTTACAGAGAGCCCATGAGAGAGCTGTGCTATCTCAGAGCAAGAGCTGCTT 1620
DB 1561 CTGTTTACAGAGAGCCCATGAGAGAGCTGTGCTATCTCAGAGCAAGAGCTGCTT 1620
QY 1621 GCGGGCTCCGCGCTCAGCTGCTGAGTGCCTGCGGCGGCACTGACATTAAGTATGCTCC 1680
DB 1621 GCGGGCTCCGCGCTCAGCTGCTGAGTGCCTGCGGCGGCACTGACATTAAGTATGCTCC 1680
QY 1681 TGTGACAGACTGTCTCTGCGCGGACCCCTATTTGGCGCTGAGAGCTCAACAGGCGC 1740
DB 1681 TGTGACAGACTGTCTCTGCGCGGACCCCTATTTGGCGCTGAGAGCTCAACAGGCGC 1740
QY 1741 TGTGCGCGTGGTGGGCTTGTGATCTTACTGATCCAGAGTGTGATGACTCGGAC 1800
DB 1741 TGTGCGCGTGGTGGGCTTGTGATCTTACTGATCCAGAGTGTGATGACTCGGAC 1800
QY 1801 ACTTCAGGCAATTCAGACCTCCGAGAGTAAATACAGTCAAGGCCCACTCCCAAAAC 1860
DB 1801 ACTTCAGGCAATTCAGACCTCCGAGAGTAAATACAGTCAAGGCCCACTCCCAAAAC 1860
QY 1861 ATCAGGCTGTGGGCGACAGACCTGTGCTGCTGACCTCTCTCAACTTGGCC 1920
DB 1861 ATCAGGCTGTGGGCGACAGACCTGTGCTGCTGACCTCTCTCAACTTGGCC 1920
QY 1921 CTG 1923

Db 712 TAAATGAGGCCCCACCACTCCATGAAGACAGTACCTGGCCTTTTGGCTCAAGCAACC 771
OY 804 TCACCTTTGAGAGCTCTGCTATGATCTGAGAGTGTGGGAGCTTTCACGGGGAGCAGCA 863
Db 772 TCACCTTTGAGAGCTCTGCTATGATCTGAGAGTGTGGGAGCTTTCACGGGGAGCAGCA 831
OY 864 CAAGGTCTACTCTCTCTTTCAGAGGAGCGGAGCTGGAGTCCGACTGCTATGCGGAGCAGT 923
Db 832 CAAGGTCTACTCTCTCTTTCAGAGGAGCGGAGCTGGAGTCCGACTGCTATGCGGAGCAGT 891
OY 924 GGTGGGCTGCTGGGCCCCGCTGCTGCAAGGGGAGATATGGGGGGGGCCACAGACCCCTGCAGAG 983
Db 892 GGTGGGCTGCTGGGCCCCGCTGCTGCAAGGGGAGATATGGGGGGGGCCACAGACCCCTGCAGAG 951
OY 984 GAAGTGGACCACTCTCTGCAAGGGGCGGCTGGAGCTGCTGCGCCCAACTGGAGCCTTA 1043
Db 952 GAAGTGGACCACTCTCTGCAAGGGGCGGCTGGAGCTGCTGCGCCCAACTGGAGCCTTA 1011
OY 1044 CTTCACACCTGCTGAGGCGATGCAACACCTGCGAGACACCTCTGCGACCAACACCTCTT 1103
Db 1012 CTTCACACCTGCTGAGGCGATGCAACACCTGCGAGACACCTCTGCGACCAACACCTCTT 1071
OY 1104 CTTGGGGTTTTCACACAGAGGGGGTGACATGTACTGTCGGCCCATCTGTGAGTACCA 1163
Db 1072 CTTGGGGTTTTCACACAGAGGGGGTGACATGTACTGTCGGCCCATCTGTGAGTACCA 1131
OY 1164 GTTGAAGAGATCCAGCGGGTGTGAGGGCCCTATAGAGATGACATGAGAAAGCCCA 1223
Db 1132 GTTGAAGAGATCCAGCGGGTGTGAGGGCCCTATAGAGATGACATGAGAAAGCCCA 1191
OY 1224 GAAGTGGAGCCGCTACACTGACCTGTACCCAGCCCTCGGGCTGGCTCTGATTAACAA 1283
Db 1192 GAAGTGGAGCCGCTACACTGACCTGTACCCAGCCCTCGGGCTGGCTCTGATTAACAA 1251
OY 1284 CTGGCATCGGGGCGGAGGCTACACAGCTCTGCGAGCTGACCCAGCAACATCCCACTT 1343
Db 1252 CTGGCATCGGGGCGGAGGCTACACAGCTCTGCGAGCTGACCCAGCAACATCCCACTT 1311
OY 1344 CGTCAAGAGAGACCCGCTGATGAGAGAGAGTGGGGCTCTGCTGAGAGCCGCCCTCT 1403
Db 1312 CGTCAAGAGAGACCCGCTGATGAGAGAGAGTGGGGCTCTGCTGAGAGCCGCCCTCT 1371
OY 1404 CGTGAAGAGAGGACCAACTTACCCACCTGCTGGCGACCGGGTTACAGAGCTTGATGG 1463
Db 1372 CGTGAAGAGAGGACCAACTTACCCACCTGCTGGCGACCGGGTTACAGAGCTTGATGG 1431
OY 1464 AGCAGCTATACAGTGTGCTCATTTGGACAGAGAGAGAGGCGTGCACAAGGCTGTGAG 1523
Db 1432 AGCAGCTATACAGTGTGCTCATTTGGACAGAGAGAGAGGCGTGCCTCAAGGCTGTGAG 1491
OY 1524 CTTGGGGCCCTGGGCTTACCTGATTTGAGAGAGCTGACCTGTTTGAACAGAGAGCCATGAG 1583
Db 1492 CTTGGGGCCCTGGGCTTACCTGATTTGAGAGAGCTGACCTGTTTGAACAGAGAGCCATGAG 1551
OY 1584 AAGCTGTGCTATCTCAGAGCAAGAAAGCTGCTCTTTGGCGGGCTCCGCTCTCAGCTGGT 1643
Db 1552 AAGCTGTGCTATCTCAGAGCAAGAAAGCTGCTCTTTGGCGGGCTCCGCTCTCAGCTGGT 1611
OY 1644 GCAGCTGCCGCTGGGCGAGCTGCAATTAAGTATGCTCCGTTGCGACAGAGTGTGCTCGGGCC 1703
Db 1612 GCAGCTGCCGCTGGGCGAGCTGCAATTAAGTATGCTCCGTTGCGACAGAGTGTGCTCGGGCC 1671
OY 1704 GGACCCCTATTTGCGCTGAGAGCTGAACACAGCCGCTGTGGCCGCTGGTGGGCGCACTT 1763
Db 1672 GGACCCCTATTTGCGCTGAGAGCTGAACACAGCCGCTGTGGCCGCTGGTGGGCGCACTT 1731
OY 1764 TGGATCTTTTACTGATCCAGCATGTGATGACTGGGACACTTCAGGCAATTTCCAACTCCG 1823
Db 1732 TGGATCTTTTACTGATCCAGCATGTGATGACTGGGACACTTCAGGCAATTTCCAACTCCG 1791
OY 1824 TGGAGTATAGATACAGTACAGGCGCATTCGCAAAAAACATCAAGGTGGGGGGGCGACAGA 1883
Db 1792 TGGAGTATAG--AAAGTACAGGCGCACT--CCCAAAAAACATCAAGGTGGGGGGGCGACAGA 1848

OY 1884 CTTGCTGCTGCGCTTGCACACCTCTCTCCAACTTGGCCCTGCGGATCCAAACCCGAGG 19442
Db 1849 CTTGCTGCTGCGCTTGCACACCTCTCTCCAACTTGGCCGATCCAGGACCTTTGGGG 1907
RESULT 5
AAD28947
ID AAD28947 standard; cDNA; 3868 bp.
XX
AC AAD28947;
XX
DT 07-MAY-2002 (first entry)
XX
DE Human MOL5a cDNA.
XX
KW Secreted molecule; MOL5a protein; MOLX; cardiomyopathy; atherosclerosis;
KW diabetes; chromosomal disorder; albinism; anaemia; psoriasis; scarring;
KW liver cirrhosis; neurodegeneration; osteoarthritis; organ rejection;
KW cerebral thrombosis; hypertension; systemic lupus erythematosus; asthma;
KW immune disease; ischaemia; immunodeficiency; rheumatoid arthritis; ulcer;
KW HIV; human immunodeficiency virus; hepatitis; hemostatic disease; pain;
KW haematopoietic; thrombolytic; thrombocytopenia; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; muscular disease; stress;
KW ocular disease; growth disorder; depression; epilepsy; contraceptive;
KW vulnery; osteopathic; haemostatic; tranquiliser; antidepressant;
KW analgesic; vasotropic; hypotensive; gene therapy; chromosome 2; ss.
XX
OS Homo sapiens.
XX
FH Location/Qualifiers
FT 1..452
FT 5'UTR
FT /*tag- a
FT 453..2954
FT CDS
FT /*tag- b
FT /product- "Human MOL5a protein"
FT /transl_except- (pos:1191..1193, aa:asp)
FT 2955..3868
FT /*tag- c
XX
PN WO200206339-A2.
XX
PD 24-JAN-2002.
XX
PE 03-JUL-2001; 2001WO-US21249.
XX
PR 03-JUL-2000; 2000US-215854P.
PR 03-JUL-2000; 2000US-215856P.
PR 03-JUL-2000; 2000US-215902P.
PR 07-JUL-2000; 2000US-216585P.
PR 07-JUL-2000; 2000US-216586P.
PR 07-JUL-2000; 2000US-216722P.
PR 17-JUL-2000; 2000US-216922P.
PR 17-JUL-2000; 2000US-216922P.
PR 27-JUL-2000; 2000US-221285P.
PR 14-FEB-2001; 2001US-268734P.
PA (CURA-) CURAGEN CORP.
XX
PI Spaderma SK, Tchernev V, Liu X, Shenoy S, Splyek K, Zethusen B;
PI Paturajan M, Taupier RJ, Rastelli L, Grose NM, Szekeres ES;
PI Alsbrook J, Lepley DM, Shen L, Burgess CE, Shinkels RA;
PI Padigaru M;
XX
DR WPI: 2002-155038/20.
XX P-PSDB: MAE18213.
XX
PT Nucleic acids encoding secreted polypeptides, designated MOLX
PT polypeptides, useful for treating a MOLX-associated disorder, e.g.,
PT cardiomyopathy, atherosclerosis, diabetes and metabolic disorders -
XX
PS Claim 3; Page 59-60; 223pp; English.
XX

QY 1863 CAGGTCGTGGCGGACACACCTGTCGTCCTGCGACCTCTCTCCCACTTGCCCT 1922
 DB 2147 CAGGTCGTGGCGGACACACCTGTCGTCCTGCGACCTCTCTCCCACTTGCCCA 2206
 OY 1923 GCCGACTCCAGCCCGAGG 1942
 DB 2207 TGGCCGCTGGACCTTTGGCG 2226

RESULT 6
 AAD28949
 ID AAD28949 standard; cDNA: 3112 BP.
 AC AAD28949;
 DT 07-MAY-2002 (first entry)
 XX Human MO15c cDNA.
 DE
 XX Secreted molecule; MO15c protein; MO1X; cardiomyopathy; atherosclerosis;
 KW diabetes; chromosomal disorder; albinism; anaemia; psoriasis; scarring;
 KW liver cirrhosis; neurodegeneration; osteoarthritis; organ rejection;
 KW cerebral thrombosis; hypertension; systemic lupus erythematosus; asthma;
 KW immune disease; ischaemia; immunodeficiency; rheumatoid arthritis; ulcer;
 KW HIV; human immunodeficiency virus; hepatitis; haemostatic disease; pain;
 KW haematopoietic; thrombolytic; thrombocytopenia; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; muscular disease; stress;
 KW ocular disease; growth disorder; depression; epilepsy; contraceptive;
 KW vulvovaginal; osteopathic; haemostatic; tranquilliser; antidepressant;
 KW analgesic; vasotropic; hypotensive; gene therapy; chromosome 2; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT 5'UTR 1..103
 FT CDS 104..2605
 FT /*tag- a
 FT /*tag- b
 FT /*product- "Human MO15c protein"
 FT sig_peptide 104..163
 FT /*tag- c
 FT mat_peptide 164..2602
 FT /*tag- d
 FT /*product- "Human MO15c mature protein"
 FT 3'UTR 2606..3112
 FT /*tag- e
 XX
 XX MO200206339-A2.
 XX
 XX 24-JAN-2002.
 XX
 XX 03-JUL-2001; 2001MO-US21249.
 XX
 XX 03-JUL-2000; 2000US-215854P.
 XX 03-JUL-2000; 2000US-215856P.
 XX 03-JUL-2000; 2000US-215902P.
 XX 07-JUL-2000; 2000US-216585P.
 XX 07-JUL-2000; 2000US-216586P.
 XX 07-JUL-2000; 2000US-216722P.
 XX 17-JUL-2000; 2000US-218622P.
 XX 17-JUL-2000; 2000US-218922P.
 XX 27-JUL-2000; 2000US-221285P.
 XX 14-FEB-2001; 2001US-268734P.
 XX
 XX (CUBA-) CUBAGEN CORP.
 XX
 XX Spederna SK, Tchervnev V, Liu X, Shenoy S, Spytek K, Zerhusen B;
 XX Pasturajan M, Taupier RJ, Rastelli L, Grosse WM, Szekeres ES;
 XX Alsobrook J, Lapeley DM, Shen L, Burgess CE, Shinkels RA;
 XX Padigaru M;
 XX WPI, 2002-155038/20.
 XX P-PSDB, AAE18215.

XX Nucleic acids encoding secreted polypeptides, designated MO1X
 PT polypeptides, useful for treating a MO1X-associated disorder, e.g.
 PT cardiomyopathy, atherosclerosis, diabetes and metabolic disorders -
 XX
 PS Claim 3; Page 63-64; 223pp; English.
 XX
 CC The patent discloses nucleic acid sequences encoding novel secreted
 CC molecule (MO1) polypeptides, designated MO1X polypeptides (i.e. a MO1
 CC protein where X is an integer from 1 to 8). Sequences of the invention
 CC are useful for treating or preventing a MO1X-associated disorder in
 CC humans. They are useful for treating or preventing cardiomyopathy,
 CC atherosclerosis and disorders related to cell signal processing and
 CC metabolic pathway modulation. The MO1X antibodies are useful for
 CC treating or preventing diabetes and disorders related to cell signal
 CC processing and metabolic pathway modulation. MO1X sequences are useful
 CC for the treatment or diagnosis of other MO1X-associated disorders, e.g.
 CC chromosomal disorders, albinism, anaemia, liver cirrhosis, psoriasis,
 CC scarring, neurodegeneration, osteoarthritis, organ rejection, cerebral
 CC thrombosis, ischaemia, hypertension, systemic lupus erythematosus, immune
 CC diseases, immunodeficiency, HIV (human immunodeficiency virus), viral,
 CC bacterial and fungal infections, hepatitis, rheumatoid arthritis, asthma,
 CC haematopoietic, thrombolytic, haemostatic diseases, thrombocytopenia,
 CC ulcers, Alzheimer's disease, Parkinson's disease, Huntington's disease,
 CC stress, depression, pain and epilepsy. They are useful for preventing
 CC chemotherapy side effects and as contraceptives. Sequences of the
 CC invention are also useful for gene therapy. The present sequence
 CC is a cDNA encoding human semaphorin 4C-like protein, MO15c. MO15c
 CC gene is localised on chromosome 2.
 XX
 XX Sequence 3112 BP; 559 A; 948 C; 957 G; 648 T; 0 other;
 S0

Query Match 75.4%; Score 1722; DB 24; Length 3112;
 Best Local Similarity 97.4%; Pred. No. 0;
 Matches 1772; Conservative 0; Mismatches 45; Indels 3; Gaps 2;

OY 123 CAGCAGGACAGCTGGCTGAAGCTCAGAGCCGGGCGTGGCCATGCGCCACACTGGC 182
 DB 61 CAGCAGGACAGCTGGCTGAAGCTCAGAGCCGGGCGTGGCCATGCGCCACACTGGC 120
 OY 183 TGTCTGGCTGCTGGCAGCAGAGCTGTGGGCTTGCGCATTTGGGCTGAGTGTGTGAA 242
 DB 121 TGTCTGGCTGCTGGCAGCAGAGCTGTGGGCTTGCGCATTTGGGCTGAGTGTGTGAA 180
 OY 243 CCTGTGGCGCGTGAAGAGTGTCTTGGGAGCTGGCCAGGTAAGAGGGGTTCTC 302
 DB 181 CCTGTGGCGCGTGAAGAGTGTCTTGGGAGCTGGCCAGGTAAGAGGGGTTCTC 240
 OY 303 CCAGACGGGCAWCCAGCACTTCTGACACTGACCTGACGAGGCCACTGGCTTCTGTA 362
 DB 241 CCAGACGGGCAWCCAGCACTTCTGACACTGACCTGACGAGGCCACTGGCTTCTGTA 300
 OY 363 CGTGGGCGCCGAGAGGCGCTGTTGCCCTTCAGCATGAGGCGCTTGACGATCAAGAGC 422
 DB 301 CGTGGGCGCCGAGAGGCGCTGTTGCCCTTCAGCATGAGGCGCTTGACGATCAAGAGC 360
 OY 423 GATCTCCGCGGAGGCGCCCGTGGAGAAAGTGAAGTATTCAGAAAGGAAAGCAAA 482
 DB 361 GATCTCCGCGGAGGCGCCCGTGGAGAAAGTGAAGTATTCAGAAAGGAAAGCAAA 420
 OY 483 CCAGACGAGAGTTCATTCATCCGCTTCTGACAGCCCTACATAGCCTCCCACTGTA 542
 DB 421 CCAGACGAGAGTTCATTCATCCGCTTCTGACAGCCCTACATAGCCTCCCACTGTA 480
 OY 543 CGTCTGTGGCACTTACGCTTTCAGACCAAGTGCACCTAGCTCAACATGCTACCTTAC 602
 DB 481 CGTCTGTGGCACTTACGCTTTCAGACCAAGTGCACCTAGCTCAACATGCTACCTTAC 540
 OY 603 TTGGAGCATGAGAGTGTGAAGTGGGAAGGGAAGTGTCCATATGACCCAGCTAAGG 662
 DB 541 TTGGAGCATGAGAGTGTGAAGTGGGAAGGGAAGTGTCCATATGACCCAGCTAAGG 600

```

OY 663 CCATCTGGCCCTTCTTGGATGATGAGCTACTCGGCCACATCAACACTTCCTGGG 722
    |||
DB 601 CCATCTGGCCCTTCTTGGATGATGAGCTACTCGGCCACATCAACACTTCCTGGG 660
OY 723 CACGGACCATATTCCTGGATGATGAGCTACTCGGCCACATCAACACTTCCTGGG 782
    |||
DB 661 CACGGACCATATTCCTGGATGATGAGCTACTCGGCCACATCAACACTTCCTGGG 720
OY 783 GGCCTTTTGGCTCAACGAACTCTACTTGTAGGCTCTGCTATGTAAGTGTGGG 842
    |||
DB 721 GGCCTTTTGGCTCAACGAACTCTACTTGTAGGCTCTGCTATGTAAGTGTGGG 780
OY 843 CAGCTTCAGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 902
    |||
DB 781 CAGCTTCAGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
OY 903 CCACTGCTATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 962
    |||
DB 841 CCACTGCTATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
OY 963 GGCCTTCAGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1022
    |||
DB 901 GGCCTTCAGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
OY 1023 TGCCCCGAACTGGAGCTCTACTTCAACCAAGTGGAGGAGTCAACACCTGAGAGAC 1082
    |||
DB 961 TGCCCCGAACTGGAGCTCTACTTCAACCAAGTGGAGGAGTCAACACCTGAGAGAC 1020
OY 1083 CTCCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1142
    |||
DB 1021 CTCCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
OY 1143 GTGGCCATCTGTGAGTACCACTTGGAGAGATCCAGCGGCTGTTGAGGCGCCATATA 1202
    |||
DB 1081 GTGGCCATCTGTGAGTACCACTTGGAGAGATCCAGCGGCTGTTGAGGCGCCATATA 1140
OY 1203 GGAGTACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1262
    |||
DB 1141 GGAGTACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
OY 1263 GCCTGGCTGTGATTAACAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1322
    |||
DB 1201 GCCTGGCTGTGATTAACAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
OY 1323 ACCGAGACATCTCTCAACTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1382
    |||
DB 1261 ACCGAGACATCTCTCAACTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
OY 1383 TCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1442
    |||
DB 1321 TCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
OY 1443 CCGGGTTACAGAGACTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1502
    |||
DB 1381 CCGGGTTACAGAGACTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
OY 1503 CTGGCTGCTCAAGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1562
    |||
DB 1441 CTGGCTGCTCAAGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
OY 1563 GTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1622
    |||
DB 1501 GTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
OY 1623 CGGCTCCGCTCTCAAGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1682
    |||
DB 1561 CGGCTCCGCTCTCAAGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
OY 1683 TGCAGACTGTGTCTTCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1742
    |||
DB 1621 TGCAGACTGTGTCTTCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
OY 1743 TGTGGCCGTGGGTGGGCACTTGGATCTTTACTGATCCAGCATGTGATGACCTCGGACAC 1802

```

```

DB 1681 TGTGGCCGCGGGGTGGCCACTCTGATCTCTACTGATCCAGCATGTGATGATCGGACAC 1740
OY 1803 TTTCAGGACATTTTCACAACTCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1862
    |||
DB 1741 TTTCAGGACATTTTCACAACTCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1797
OY 1863 CACGGTGTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1922
    |||
DB 1798 CACGGTGTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1857
OY 1923 GCCCGACTCCAAACCCGAGAG 1942
    |||
DB 1858 TCCCGCTGGAGACTTGGGG 1877

RESULT 7
AAD28948
ID AAD28948 standard; cDNA; 2558 BP.
XX
AC AAD28948;
XX
DT 07-MAY-2002 (first entry)
XX
DE Human MOL5b cDNA.
XX
KW Secreted molecule; MOL5b protein; MOLX; cardiomyopathy; atherosclerosis;
KW diabetes; chromosomal disorder; albinism; anemia; psoriasis; scarring;
KW liver cirrhosis; neurodegeneration; osteoarthritis; organ rejection;
KW cerebral thrombosis; hypertension; systemic lupus erythematosus; asthma;
KW immune disease; ischemia; immunodeficiency; rheumatoid arthritis; ulcer;
KW HIV; human immunodeficiency virus; hepatitis; haemostatic disease; pain;
KW haemopoietic; thrombolytic; thrombocytopenia; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; muscular disease; stress;
KW ocular disease; growth disorder; depression; epilepsy; contraceptive;
KW vulvectomy; osteoporosis; haemostatic; tranquilizer; antidepressant;
KW analgesic; vasodilator; hypotensive; gene therapy; chromosome 2; 99.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT 5'UTR 1..20
FT CDS 21..2522
FT FT /*tag= "Human MOL5b protein"
FT FT /*tag= "Human MOL5b protein"
FT FT 2523..2558
FT FT /*tag= "c"
XX
XX MO200206339-A2.
XX
XX 24-JAN-2002.
XX
XX 03-JUL-2001: 2001WO-US21249.
XX
XX 03-JUL-2000: 2000US-215854P.
XX 03-JUL-2000: 2000US-215856P.
XX 03-JUL-2000: 2000US-215902P.
XX 07-JUL-2000: 2000US-216585P.
XX 07-JUL-2000: 2000US-216586P.
XX 07-JUL-2000: 2000US-216722P.
XX 17-JUL-2000: 2000US-218622P.
XX 17-JUL-2000: 2000US-218992P.
XX 27-JUL-2000: 2000US-221285P.
XX 14-FEB-2001: 2001US-268734P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Spaderne SK, Tchiernev V, Liu X, Shenoy S, Spytek K, Zerhusen B;
XX Patuturajan M, Traupler RJ, Rastelli L, Grosse WM, Szekeres ES;
XX Alsbrook J, Lepley DM, Shen L, Burgess CE, Shinkets RA;
XX Padigaru M;
XX

```

DR WPI: 2002-155038/20.
DR P-PSDB: AAE18214.
XX Nucleic acids encoding secreted polypeptides, designated MOIX
PT polypeptides, useful for treating a MOIX-associated disorder, e.g.
PT cardiomyopathy, atherosclerosis, diabetes and metabolic disorders -
XX claim 3: Page 61; 223pp; English.
XX
CC The patent discloses nucleic acid sequences encoding novel secreted
CC molecule (MOI) polypeptides, designated MOIX polypeptides (i.e. a MOI
CC protein where X is an integer from 1 to 8). Sequences of the invention
CC are useful for treating or preventing a MOIX-associated disorder in
CC humans. They are useful for treating or preventing cardiomyopathy,
CC atherosclerosis and disorders related to cell signal processing and
CC metabolic pathway modulation. The MOIX antibodies are useful for
CC treating or preventing diabetes and disorders related to cell signal
CC processing and metabolic pathway modulation. MOIX sequences are useful
CC for the treatment or diagnosis of other MOIX-associated disorders, e.g.
CC chromosomal disorders, albinism, anaemia, liver cirrhosis, psoriasis,
CC scarring, neurodegeneration, osteoarthritis, organ rejection, cerebral
CC thrombosis, ischaemia, hypertension, systemic lupus erythematosus, immune
CC diseases, immunodeficiency, HIV (human immunodeficiency virus), viral,
CC bacterial and fungal infections, hepatitis, rheumatoid arthritis, asthma,
CC haematopoietic, thrombolytic, haemostatic diseases, thrombocytopenia,
CC ulcers, Alzheimer's disease, Parkinson's disease, Huntington's disease,
CC ocular disease, muscular diseases, growth disorders, loss of libido,
CC stress, depression, pain and epilepsy. They are useful for preventing
CC chemotherapy side effects and as contraceptives. Sequences of the
CC invention are also useful for gene therapy. The present sequence
CC is a cDNA encoding human semaphorin 4C-like protein, MOI5b. MOI5b
CC gene is localised on chromosome 2.
XX
XX Sequence 2558 bp; 468 A; 788 C; 801 G; 501 T; 0 other;
Query Match 74.4%; Score 1699; DB 24; Length 2558;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1749; Conservative 0; Mismatches 45; Indels 3; Gaps 2;

DB 481 ATGGGAAGGCAAGTGTCCCTATGACCCAGCTAAGGCCATGCTGCGCTTCTTGATG 540
QY 686 GTGAGCTGTACTGGGCCACACTCAACAACCTTCTGGGACGGAACCCATTATCCGCTGA 745
DB 541 GTGAGCTGTACTGGGCCACACTCAACAACCTTCTGGGACGGAACCCATTATCCGCTGA 600
QY 746 ACATGGGGCCCAACCACTCATGAAGACAGATGAGTGGGCTTTGGCTCAAGCAACCTC 805
DB 601 ACATGGGGCCCAACCACTCATGAAGACAGATGAGTGGGCTTTGGCTCAAGCAACCTC 660
QY 806 ACTTTGAGGCTCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 865
DB 661 ACTTTGAGGCTCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 720
QY 866 AGGCTCTCTTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 925
DB 721 AGGCTCTCTTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
QY 926 TGGCTCGTGTGGCCGCTGTGTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 985
DB 781 TGGCTCGTGTGGCCGCTGTGTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
QY 986 AGTGACACAGTCTGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1045
DB 841 AGTGACACAGTCTGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
QY 1046 TCAACAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1105
DB 901 TCAACAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
QY 1106 TTGGGGTTTTTCAAGACACAGTGGGGTGAATGATGATGATGATGATGATGATGATG 1165
DB 961 TTGGGGTTTTTCAAGACACAGTGGGGTGAATGATGATGATGATGATGATGATGATG 1020
QY 1166 TGAAGAGATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1225
DB 1021 TGAAGAGATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
QY 1226 AGTGAGACCGTACACTGACCTTATACCAAGCCCTGGGAGGAGGAGGAGGAGGAGGAG 1285
DB 1081 AGTGAGACCGTACACTGACCTTATACCAAGCCCTGGGAGGAGGAGGAGGAGGAGGAG 1140
QY 1286 GGCATGGGCGCCACAGGCTACACAGCTCCCTGAGCTACCCGACACATCTCAACTTG 1345
DB 1141 GGCATGGGCGCCACAGGCTACACAGCTCCCTGAGCTACCCGACACATCTCAACTTG 1200
QY 1346 TCAAGGAACACCCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1405
DB 1201 TCAAGGAACACCCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1406 TGAAGAGAGGACCAACTTCAACCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1465
DB 1261 TGAAGAGAGGACCAACTTCAACCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
QY 1466 CCACCTATACAGTCTCTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1525
DB 1321 CCACCTATACAGTCTCTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1526 TGGGGGCTGGGCTTACCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1585
DB 1381 TGGGGGCTGGGCTTACCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1586 GCGTGTGTATATCAAG 1645
DB 1441 GCGTGTGTATATCAAG 1500
QY 1646 AGCTGCCGTGGCCGAGTGCATTAAGATATGCTCTGTGACAGACTGTGTCTGCGCGG 1705
DB 1501 AGCTGCCGTGGCCGAGTGCATTAAGATATGCTCTGTGACAGACTGTGTCTGCGCGG 1560
QY 1706 ACCCTATGCGCCTGTGAGGCTCAACACAGCCGCTGTGTGAGGAGGAGGAGGAGGAG 1765
DB 1561 ACCCTATGCGCCTGTGAGGCTCAACACAGCCGCTGTGTGAGGAGGAGGAGGAGGAG 1620

OY 1766 GATCTTACTGATCCAGCATGTGATGACCTCGGACATTCGACATTTTCGCACTCCGTTG 1825
|||||
DB 1621 GATCTCTACTGATCCAGCATGTGATGACCTCGGACATTCGACATTCGCACTCCGTTG 1680
OY 1826 GAGATAGATTAAGTCAGGCGCCACATNCCAAAATCAACGCTGGTGGGGGACAGAC 1885
|||||
DB 1681 GAGATAGG--AAAGTCAGGCGCCACT--CCCAAAAACATCAGGTGGTGGGGGACAGAC 1737
OY 1886 TGGTCTGCTGCTGCGACCTCTCTCCAACTTGGCCCTGGCGACATCCAAACCCGAGG 1942
|||||
DB 1738 TGGTCTGCTGCTGCGACCTCTCTCCAACTTGGCCCTGGCGACATTCGAGCTTTGGGG 1794

RESULT 8
ABN83983
ID ABN83983 standard; DNA; 3776 BP.
XX
AC ABN83983;
XX
DT 06-SEP-2002 (first entry)
XX
DE Human gene sequence #30.
XX
KW Human; brain; tonsil; hippocampus; foetal brain; diagnosis; gene; ds.
XX
OS Homo sapiens.
XX
FII Key Location/Qualifiers
FT CDS 232..2692
FT /tag= a
XX
PN WO200252005-A1.
XX
PD 04-JUL-2002.
XX
PF 20-DEC-2001; 2001WO-JP11217.
XX
PR 22-DEC-2000; 2000JP-0389742.
XX
PA (KAZU-) KAZUSA DNA RES INST FOUND.
XX (CELE-) CELESTAR LEXICO-SCI LTD.
XX
PI Ohara O, Nagase T, Nakajima D;
XX
DR WPI: 2002-500762/53.
XX P-PSDB: ABB97963.
XX
PT Genes and their expression products cloned from human cDNA libraries
PT for treatment and diagnosis of diseases associated with their
PT expression -
XX
PS Claim 1(a): Page 119-125; 238pp; Japanese.
XX
CC The invention relates to DNA encoding polypeptides directly cloned from
CC cDNA libraries originating in adult whole brain, human tonsil, human
CC adult hippocampus and human foetal whole brain. Polypeptides and
CC polynucleotides of the invention may be used in the investigation of
CC differential expression of the DNA sequences in normal subjects and
CC disease patients. They may also be used in the production of antibodies,
CC oligonucleotide probes and DNA chips for diagnosis and identification
CC of drugs for treatment of diseases with which the DNA sequences are
CC associated. The sequences given in records ABN83954-ABN83984 represent
CC human gene sequences of the invention.
XX
SO Sequence 3776 BP; 696 A; 1088 C; 1191 G; 801 T; 0 other;

Query Match 70.8%; Score 1617.6; DB 24; Length 3776;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1552; Conservative 0; Mismatches 19; Indels 3; Gaps 2;

OY 269 CTGGGACCTGGCCGCTAGACGGGGTCTCCCAACCGGCAATCCAGAGACTTCCTGA 328
||| |||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 494 CTGCAGAGCTGGCCACGAGTGAAGCGGGTCTCCCAAGACCGGCATCCAGACTTCCTGA 553
OY 329 CACTGAGCTGAGAGAGAGCCCACTGGGCTTCTGATAGCTGGGCGCCGAGAGCCCTGTTG 388
|||||
DB 554 CACTGAGCTGAGAGAGAGCCCACTGGGCTTCTGATAGCTGGGCGCCGAGAGCCCTGTTG 613
OY 389 CCTTCAGATGAGAGGCGCTGAGAGCTGCAAGAACGATCTCTGGAGGCCCCCGGTGAGA 448
|||||
DB 614 CCTTCAGATGAGAGGCGCTGAGAGCTGCAAGAACGATCTCTGGAGGCCCCCGGTGAGA 673
OY 449 AGAAGCTAGTATATCCAGAAAGGAAAGAACACACAGAGCTTCACTTATATG 508
|||||
DB 674 AGAAGCTAGTATATCCAGAAAGGAAAGAACACACAGAGCTTCACTTATATC 733
OY 509 GCTTCCTGAGAGCCCTCAATAGCTCCACCTGATAGCTGTGGACCTAGGCTTCACG 568
|||||
DB 734 GCTTCCTGAGAGCCCTCAATAGCTCCACCTGATAGCTGTGGACCTAGGCTTCACG 793
OY 569 CCAAGTGACCTAAGCTCAACATGCTCACCCTTCACTTTGGAGCATGAGAGTTGAAGATG 628
|||||
DB 794 CCAAGTGACCTAAGCTCAACATGCTCACCCTTCACTTTGGAGCATGAGAGTTGAAGATG 853
OY 629 GGAAGGCAAGTGTCCCTATGACCCAGCTAAGGCCATGCTGCTTCTTGTGATGATG 688
|||||
DB 854 GGAAGGCAAGTGTCCCTATGACCCAGCTAAGGCCATGCTGCTTCTTGTGATGATG 913
OY 689 ACCTGATCTGAGGCAACACTCAACACTTCTGGGCAAGGACCATTAATCTGCTGATGA 748
|||||
DB 914 ACCTGATCTGAGGCAACACTCAACACTTCTGGGCAAGGACCATTAATCTGCTGATGA 973
OY 749 TGGGGCCCCCAGCTCCTATGAAAGACAGACTGCTGCTTTTGGCTCAAGAACTGACT 808
|||||
DB 974 TGGGGCCCCCAGCTCCTATGAAAGACAGACTGCTGCTTTTGGCTCAAGAACTGACT 1033
OY 809 TTGTAGGCTCTGCTATGATAGTGAAGAGTGTGGGAGCTTCAAGGGGGAGAGCAAG 868
|||||
DB 1034 TTGTAGGCTCTGCTATGATAGTGAAGAGTGTGGGAGCTTCAAGGGGGAGAGCAAG 1093
OY 869 TCTACTTCTTCTTCAAGGAGCGGCGAGTGCAGTCCGACTGATGCGGAGCAGAGTGTG 928
|||||
DB 1094 TCTACTTCTTCTTCAAGGAGCGGCGAGTGCAGTCCGACTGATGCGGAGCAGAGTGTG 1153
OY 929 CTCTGTGGCGCGCTGTCTGCAAGGGCGATATGGGGCGCGACGAGACCTTCGAGAGAGT 988
|||||
DB 1154 CTCTGTGGCGCGCTGTCTGCAAGGGCGATATGGGGCGCGACGAGACCTTCGAGAGAGT 1213
OY 989 GGACCACTGCTCTGAAGGGCGGCGGCTGCTGCTGCGGCAAGTGGCAGCTCTACTTA 1048
|||||
DB 1214 GGACCACTGCTCTGAAGGGCGGCGGCTGCTGCTGCGGCAAGTGGCAGCTCTACTTA 1273
OY 1049 ACCAGCTGAGGCGATGCAACCTTCGAGAGACCTCTGCGCAACACACACTTCTTTG 1108
|||||
DB 1274 ACCAGCTGAGGCGATGCAACCTTCGAGAGACCTCTGCGCAACACACACTTCTTTG 1333
OY 1109 GGGTTTTCACACAGTGGGTGACATATCTGTCGGCCATCTGTGAGTACAGTGTG 1168
|||||
DB 1334 GGGTTTTCACACAGTGGGTGACATATCTGTCGGCCATCTGTGAGTACAGTGTG 1393
OY 1169 AAGAGTCCAGGGGGTGTGAGGGCCCTATTAAGAGTACCATGAGAGAGCCGGAAGT 1228
|||||
DB 1394 AAGAGTCCAGGGGGTGTGAGGGCCCTATTAAGAGTACCATGAGAGAGCCGGAAGT 1453
OY 1229 GGGAGCGCTACACTGACCTGTATCCAGAGCCCTGCGGCTGCTGTCATTAACAAGTGGC 1288
|||||
DB 1454 GGGAGCGCTACACTGACCTGTATCCAGAGCCCTGCGGCTGCTGTCATTAACAAGTGGC 1513
OY 1289 ATGCGCGCAGGCTACACACAGCTCCCTGAGAGTACCCGACATCTCAACTTCGTGA 1348
|||||
DB 1514 ATGCGCGCAGGCTACACACAGCTCCCTGAGAGTACCCGACATCTCAACTTCGTGA 1573
OY 1349 AGAAGCACCCGCTATGAGAGAGAGAGAGTGGGGCTCTGGTGGAGCCGCCCTGCTGTGA 1408
|||||
DB 1574 AGAAGCACCCGCTATGAGAGAGAGAGTGGGGCTCTGGTGGAGCCGCCCTGCTGTGA 1633

QY	1409	AGAAAGGACACAACTTTCACCCACCTGGTGGCCGACCGGGTTACAGAGACTTATATGAGCCCA	1468
Db	1634	AGAAAGGACACAACTTTCACCCACCTGGTGGCCGACCGGGTTACAGAGACTTATATGAGCCCA	1633
QY	1469	CCATPACAGTGTCTGTTCATTGGGCACAGAGAGACGGCTGGCTCTCAAGCTGTGAGCCCTGG	1528
Db	1694	CCATPACAGTGTCTGTTCATTGGGCACAGAGAGAGCGCTGGCTCTCAAGCTGTGAGCCCTGG	1753
QY	1529	GGCCCTGGGTTCCACCTGATTGAGAGCTGCAGCTGTTTGACACAGAGCCCATYAGAAAGCC	1588
Db	1754	GGCCCTGGGTTCCACCTGATTGAGAGCTGCAGCTGTTTGACACAGAGCCCATYAGAAAGCC	1813
QY	1589	TGCTGCTATCTCAGAGCAGAAGAGTGGCTTTGGCCGGTCCCGCTTCAGCTGGGCGAGC	1648
Db	1814	TGCTGCTATCTCAGAGCAGAAGAGTGGCTTTGGCCGGTCCCGCTTCAGCTGGGCGAGC	1873
QY	1649	TGCCCGTGGCCGACGTGCATTAAGTATGAGTCTCTGTGCAGACTGTGTCTCGCCGGGAGCC	1708
Db	1874	TGCCCGTGGCCGACGTGCATTAAGTATGAGTCTCTGTGCAGACTGTGTCTCGCCGGGAGCC	1933
QY	1709	CCATTTGGCGCTGAGAGCTCAACACACAGCCGCTGTGTGGCCGTGAGGCGCACTTTGGAT	1768
Db	1934	CCATTTGGCGCTGAGAGCTCAACACACAGCCGCTGTGTGGCCGTGAGGCGCACTTTGGAT	1993
QY	1769	CTTACTAGATCCAGCATGTGTATGACCTGGGACACTTCAGGATTTGCAACCTTCGGTGGCA	1828
Db	1994	CTTACTAGATCCAGCATGTGTATGACCTGGGACACTTCAGGATTTGCAACCTTCGGTGGCA	2053
QY	1829	GTAAGATACACTCAGGCGCCACTTCCCAAAAACATCACGTTGGTGGCGGCACAGACCTGG	1888
Db	2054	GTAAG--AAAGTACAGGCCCACT--CCCAAAAACATCACGTTGGTGGCGGCACAGACCTGG	2110
QY	1889	TGCTGGCCCTGGCAGCTCTCTCTCACAACCTTGGCCCTTGGCAGTCCAAACCCGAGG	1942
Db	2111	TGCTGGCCCTGGCAGCTCTCTCTCACAACCTTGGCCCAATGGCCGCTGGAGACCTTGGGG	2164
RESULT 9			
ID	ABX92031	standard; cDNA: 8095 BP.	
XX	ABX92031:		
AC	ABX92031:		
XX	08-MAY-2003 (first entry)		
DT			
XX	Lung specific nucleic acid (LSNA) #73.		
DE			
XX	Human; gene therapy; vaccine; lung specific antigen; cancer diagnosis;		
KW	cancer monitoring; cancer staging; cancer imaging; lung cancer;		
KW	non-cancerous diseases of the lung; transgenic animal; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
XX	WO200268653-A2.		
PN			
PD	06-SEP-2002.		
XX			
XX	21-NOV-2001; 2001WO-US43612.		
PF			
XX	22-NOV-2000; 2000US-252500P.		
PR			
XX	(DIAD-) DIADENUS INC.		
PA			
XX			
XX	Macina RA, Reclpon H, Chen S, Sun Y, Liu C;		
PI			
XX	WPI: 2002-713376/77.		
DR			
XX			
PT	New isolated human nucleic acid molecule and polypeptide, useful for		
PT	identifying, diagnosing, monitoring, staging, imaging and treating lung		
PT	cancer and non-cancerous diseases of the lung		
XX			
PS	Claim 1; Page 221-226; 389pp; English.		

Query Match	Best Local Similarity	Score	DB	Length
Matches 1448; Conservative 0; Mismatches 30; Indels 6; Gaps 3;	97.6%; Pred. No. 0;	1394;	24;	8095;
OY	462	TATCCAGAAAGGAGAAACACACAGACGAGTCTTCAACTTCATCCGCTTCGACACC	521	
Db	5887	TAAACCGACACCTTCTTCGCCCCAGACGAGTGGTCTCAACTTCATCCGCTTCGACACC	5948	
OY	522	CTAAGAGCCGCCACCGTACGCTGAGACCTAGACCCCTTCCAGCCCAAGTCGACCTA	581	
Db	5947	CTACAATGCCCTCCACCGTACGCTGAGACCTAGACCCCTTCCAGCCCAAGTCGACCTA	6008	
OY	582	CGTCACACATGCTCACTTCACCTTGGAGACATGAGAGTTTGAAGATGGGAAGGCAATGT	641	
Db	6007	CGTCACACATGCTCACTTCACCTTGGAGACATGAGAGTTTGAAGATGGGAAGGCAATGT	6066	
OY	642	TCCCTATGACCCACCTAAGGGCCATGCTGCGCTTCTTGATGATGGTGAAGTCACTGAC	701	
Db	6067	TCCCTATGACCCACCTAAGGGCCATGCTGCGCTTCTTGATGATGGTGAAGTCACTGAC	6128	
OY	702	CACACTCAACAACTTCCCTGGGACAGGAACCCATTATCTGGGTAACTATGGGGCCCA	761	
Db	6127	CACACTCAACAACTTCCCTGGGACAGGAACCCATTATCTGGGTAACTATGGGGCCCA	6188	
OY	762	CTCCATGAAAGACAGAGTACTGAGCTTGGCTCAACACATCTACTTGTAGGCTGAC	821	
Db	6187	CTCCATGAAAGACAGAGTACTGAGCTTGGCTCAACACATCTACTTGTAGGCTGAC	6248	
OY	822	CTATGTACTCTGAGAGTGTGGGACAGCTTACAGGGGGACAGACAAAGTCTACTTCTT	881	
Db	6247	CTATGTACTCTGAGAGTGTGGGACAGCTTACAGGGGGACAGACAAAGTCTACTTCTT	6308	
OY	882	CAGGAGCGGGGCAATGGAGTCCGACTGCTATGCCGAGACAGAGTGGTGGCTCGTGGCCG	941	
Db	6307	CAGGAGCGGGGCAATGGAGTCCGACTGCTATGCCGAGAGAGTGGTGGCTCGTGGCCG	6366	
OY	942	TGTGTGCAAGAGGCGATATGGGGGGGGGACGGACCCCTGACAGAGGAAGTGGACAGT	1001	
Db	6367	TGTGTGCAAGAGGCGATATGGGGGGGGGACGGACCCCTGACAGAGGAAGTGGACAGT	6428	
OY	1002	GAAAGCGCGGCTGGCATGCTGTGCCCCGACACTGGGCAAGCTACTTCAACACAGTCGACGC	1061	
Db	6427	GAAAGCGCGGCTGGCATGCTGTGCCCCGACACTGGGCAAGCTACTTCAACACAGTCGACGC	6488	
OY	1062	GATGCACACCCCTGACAGACACTCTCTGGCACAACACACACTTCTTGGGGTTTTCGAAGC	1121	
Db	6487	GATGCACACCCCTGACAGACACTCTCTGGCACAACACACACTTCTTGGGGTTTTCGAAGC	6544	
OY	1122	ACAGTGGGGGTACATGATACCTGTGGGCCATCTGTGATGATACCAAGTTGGGAAGATCCACGC	1188	
Db	6547	ACAGTGGGGGTACATGATACCTGTGGGCCATCTGTGATGATACCAAGTTGGGAAGATCCACGC	6606	
OY	1182	GGTGTGTTAAGAGCCCTTATTAAGAGATACCATGAGGAACCCGAAAGTGGACCGCTACAC	1241	
Db	6607	GGTGTGTTAAGAGCCCTTATTAAGAGATACCATGAGGAACCCGAAAGTGGACCGCTACAC	6666	
OY	1242	TGACCTGTACCAAGCCCTGCGGCTGAGTGTGATTAACAATGGATCGGCGCACGG	1301	

Db 6667 TGACCTTACCCAGCCCTCGGCTGCTGCTGCTTAAACAATGCGATCGGCGCCACGG 6726
 Oy 1302 CTACACAGCTCCCTGAGAGTACCCGACACATCTCTCAACTCTCTCAAGAGACCCGCT 1361
 Db 6727 CTACACAGCTCCCTGAGAGTACCCGACACATCTCTCAACTCTCTCAAGAGACCCGCT 6786
 Oy 1352 GATGAGAGAGAGTGGGGGCTCGGTGAGACCGCCCTGCTCGTGAAGAGAGGACCA 1421
 Db 6787 GATGAGAGAGAGTGGGGGCTCGGTGAGACCGCCCTGCTCGTGAAGAGAGGACCA 6846
 Oy 1422 CTTCACCCAGCTGTGGCGGACCGGTTACAGAGCTTATGAGAGCCACTATACAGTGT 1481
 Db 6847 CTTCACCCAGCTGTGGCGGACCGGTTACAGAGCTTATGAGAGCCACTATACAGTGT 6906
 Oy 1482 GTTCATTTGAGCAGAGAGCGCTGCTCTCAAGGCTTGAAGCTTGGGCGCTGGGTTCA 1541
 Db 6907 GTTCATTTGAGCAGAGAGCGCTGCTCTCAAGGCTTGAAGCTTGGGCGCTGGGTTCA 6966
 Oy 1542 CCTGATTTAGAGAGCTGACAGCTGTGAGACGAGAGCCCATGAGAACTGGTGTATCTCA 1601
 Db 6967 CCTGATTTAGAGAGCTGACAGCTGTGAGACGAGAGCCCATGAGAACTGGTGTATCTCA 7026
 Oy 1602 GAGCAAG---AAGCTGCTCTTGGCCGCTCCGCTCTCAGCTGTGAGAGCCCGTGGC 1658
 Db 7027 GAGCAAGGTAAAGCTGCTCTTGGCCGCTCCGCTCTCAGCTGTGAGAGCCCGTGGC 7086
 Oy 1659 CGAGTGCATTAAGTATGCTCTGCTGTGAGAGCTGTGTCTCTCGCCGCGGAGACCCCTATTGGCG 1718
 Db 7087 CGAGTGCATTAAGTATGCTCTGCTGTGAGAGCTGTGTCTCTCGCCGCGGAGACCCCTATTGGCG 7146
 Oy 1719 CTGGAGCTTCAACACAGCCGCTGTGAGAGCTGTGTGAGAGCTGTGTGAGAGCTGTGTGAG 1778
 Db 7147 CTGGAGCTTCAACACAGCCGCTGTGAGAGCTGTGTGAGAGCTGTGTGAGAGCTGTGTGAG 7206
 Oy 1779 CGAGTGCATTAAGTATGCTCTGCTGTGAGAGCTGTGTGAGAGCTGTGTGAGAGCTGTGTGAG 1838
 Db 7207 CGAGTGCATTAAGTATGCTCTGCTGTGAGAGCTGTGTGAGAGCTGTGTGAGAGCTGTGTGAG 7264
 Oy 1839 GTGAGAGCCACTNCCCAAAACATACAGGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1898
 Db 7265 GTGAGAGCCACTNCCCAAAACATACAGGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 7323
 Oy 1899 CGACCTCTCTTCAACTTGGCCCTGCGGAGCTTCCAAAGCCCGAGG 1942
 Db 7324 CGACCTCTCTTCAACTTGGCCCTGCGGAGCTTCCAAAGCCCGAGG 7367
 RESULT 10
 ABT13397
 ID ABT13397 standard; DNA: 8144 BP.
 XX
 AC ABT13397;
 XX
 DT 30-JAN-2003 (first entry)
 XX
 DE Breast specific related polynucleotide SEQ ID No 112.
 XX
 KW Cytostatic; BSP-agonist; BSP-antagonist; vaccine; gene therapy; cancer;
 KM metastatic; breast cancer; breast specific; human; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200277232-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-NOV-2001; 2001WO-US43815.
 XX
 PR 22-NOV-2000; 2000US-252509P.
 XX
 PA (DIAD-) DIADEXUS INC.
 XX
 PI Saleeda S, MacIna RA, Reclon H, Pluta J, Sun Y, Liu C;

XX
 DR WPI: 2003-018927/01.
 XX
 PT New isolated nucleic acid molecule, useful for treating breast cancer,
 PT and diagnosing or monitoring the presence of metastases of breast
 PT cancer in a patient
 PS Claim 1: Page 256-260; 377pp; English.
 XX
 CC The invention relates to a novel isolated nucleic acid molecule
 CC comprising: a sequence encoding a sequence comprising 11-1518 amino
 CC acids; a sequence comprising 190-8144 bp; or a sequence that selectively
 CC hybridizes to, or having at least 60% identity with the 11-1518 amino
 CC acid or 190-8144 nt sequence. The polypeptide and the nucleic acid are
 CC useful for treating breast cancer, and diagnosing or monitoring the
 CC presence of metastases of breast cancer in a patient. The polynucleotides
 CC of the invention can be used to treat disorders by gene therapy. This
 CC polynucleotide represents a breast specific related sequence of the
 CC invention.
 XX
 SQ Sequence 8144 BP; 1900 A; 2247 C; 2053 G; 1944 T; 0 other;
 XX
 Query Match 61.0%; Score 1394; DB 25; Length 8144;
 Best Local Similarity 97.6%; Pred. No. 0;
 Matches 1448; Conservative 0; Mismatches 30; Indels 6; Gaps 3;
 Oy 462 TATCCAGAAAGGAAAGAACACACAGAGAGTGTCAACTTATCCGCTTCTCAGAC 521
 Db 5936 TAAACGAGACCTTCTGCGCCACAGAGAGTGTCAACTTATCCGCTTCTCAGAC 5995
 Oy 522 CTACAAATGCTCCACCTGTAGTGTGAGAGCTTACAGCTTCCAGCCCAATGTGACCTA 581
 Db 5996 CTACAAATGCTCCACCTGTAGTGTGAGAGCTTACAGCTTCCAGCCCAATGTGACCTA 6055
 Oy 582 CTTCAACATGCTCACCTTCACTTGTGAGAGAGAGAGTGTGAAGATGGGAAGCAAGT 641
 Db 6056 CTTCAACATGCTCACCTTCACTTGTGAGAGAGAGAGTGTGAAGATGGGAAGCAAGT 6115
 Oy 642 TCCCTATGACCCAGCTAAGAGGAGCTGAGCTTCTTGTGATGTGAGAGCTGAGCTGAG 701
 Db 6116 TCCCTATGACCCAGCTAAGAGGAGCTGAGCTTCTTGTGATGTGAGAGCTGAGCTGAG 6175
 Oy 702 CACACTCAACAACTTCTGAG 761
 Db 6116 CACACTCAACAACTTCTGAG 6235
 Oy 762 CTTCATGAAG 821
 Db 6236 CTTCATGAAG 6295
 Oy 822 CTATGTACCTGAG 881
 Db 6296 CTATGTACCTGAG 6355
 Oy 882 CAGGAG 941
 Db 6356 CAGGAG 6415
 Oy 942 TGTCTGCAAG 1001
 Db 6416 TGTCTGCAAG 6475
 Oy 1002 GAAGCGCGCGGCTGAG 1061
 Db 6476 GAAGCGCGCGGCTGAG 6535
 Oy 1062 GATGACACACCTGAG 1121
 Db 6536 GATGACACACCTGAG 6595
 Oy 1122 ACAGTGGGAG 1181
 Db 6596 ACAGTGGGAG 6655


```
OY 1182 GGTGTTGAGGGCCCTATTAAGAGTACATGAGGAAGCCAGAAAGTGGAGCCGTACAC 1241
|||||
Db 6656 GGTGTTGAGGGCCCTATTAAGAGTACATGAGGAAGCCAGAAAGTGGAGCCGTACAC 6715
OY 1242 TGACCCGTATACCAAGCCCTGGCCCTGCTGCTGATTAACAACTGGATGGCCACGG 1301
|||||
Db 6716 TGACCCGTATACCAAGCCCTGGCCCTGCTGCTGATTAACAACTGGATGGCCACGG 6775
OY 1302 CTACACAGCTCCCTGGAGCTACCCGCAACATCCCACTTGCTCAAGAGCCCGCT 1361
|||||
Db 6776 CTACACAGCTCCCTGGAGCTACCCGCAACATCCCACTTGCTCAAGAGCCCGCT 6835
OY 1362 GATGAGAGAGCAGGTGGGGCTGCTGAGAGCCGCCCCCTGCTGTAAGAGGACACAA 1421
|||||
Db 6836 GATGAGAGAGCAGGTGGGGCTGCTGAGAGCCGCCCCCTGCTGTAAGAGGACACAA 6895
OY 1422 CTTCACCCACCTGCTGGCCGACCGGGTTACAAGACTTGATGAGACCACTATACAGTCT 1481
|||||
Db 6896 CTTCACCCACCTGCTGGCCGACCGGGTTACAAGACTTGATGAGACCACTATACAGTCT 6955
OY 1482 GTTCATTGGACAGAGAGCGCTGGCTGCTCAAGGCTGTAGGCTTGCGCCCTGGGTTCA 1541
|||||
Db 6956 GTTCATTGGACAGAGAGCGCTGGCTGCTCAAGGCTGTAGGCTTGCGCCCTGGGTTCA 7015
OY 1542 CCTGATTGAGAGAGCTGCACTGTTTGAACAGAGCCCATGAGAAGCCTGGTGTATCTCA 1601
|||||
Db 7016 CCTGATTGAGAGAGCTGCACTGTTTGAACAGAGCCCATGAGAAGCCTGGTGTATCTCA 7075
OY 1602 GAGCAGG---AAGTGTCTTTTCCGGCTCCCGCTCTACAGCTGTGCAAGCTGCCCTGCGC 1658
|||||
Db 7076 GAGCAGGTTAAAGCTGCTCTTTCGCGCTCCCGCTCTACAGCTGTGCAAGCTGCCCTGCGC 7135
OY 1659 CGACTGATTAAGATGCTGCTGCTGTGCAAGCTGTGCTGCGCCGGAGACCCCTATTGCGC 1718
|||||
Db 7136 CGACTGATTAAGATGCTGCTGCTGTGCAAGCTGTGCTGCGCCGGAGACCCCTATTGCGC 7195
OY 1719 CTGAGCGCTACAGACAGCGCTGTGTGGCGCTGGTGGCCACTTTGATCTTACTGAT 1778
|||||
Db 7196 CTGAGCGCTACAGACAGCGCTGTGTGGCGCTGGTGGCCACTTTGATCTTACTGAT 7255
OY 1779 CAGCAGATGATGACCTCGGACACTTACAGCATTTGCAACCTCGGTGGCAGTAAGATACA 1838
|||||
Db 7256 CAGCAGATGATGACCTCGGACACTTACAGCATTTGCAACCTCGGTGGCAGTAAG--AAA 7313
OY 1839 GTAGAGCCCACTACCAAAACATCAGAGTGTGGGCGGACAGACCTGTGCTGCCCTG 1898
|||||
Db 7314 GTAGAGCCCACT--CCCAAAACATCAGAGTGTGGGCGGACAGACCTGTGCTGCCCTG 7372
OY 1899 CCACTCTCTCCCAACTTGGCCCTGCGGCACTCCAAACCCGAGG 1942
|||||
Db 7373 CCACTCTCTCCCAACTTGGCCCTGCGGCACTCCAAACCCGAGG 7416

RESULT 11
AAF29461
ID AAF29461 standard; cdna; 3503 BP.
XX
XX AAF29461;
XX
XX 03-APR-2001 (first entry)
XX
XX Murine M-Sema-F cdna.
XX
XX Mouse; M-Sema-F; INTERCEPT 217; INTERCEPT 297; TANGO 276;
XX TANGO 292; TANGO 325; TANGO 331; TANGO 332; cytostatic; antiinflammatory;
XX antiarhythmic; antiapoptotic; gene therapy; cancer;
XX inflammatory disorder; cardiac disorder; arrhythmia; skin disorder;
XX psoriasis; ss.
XX
XX Mus sp.
XX
XX W0200100638-A2.
```

```
XX 04-JAN-2001.
PD
XX
XX 16-JUN-2000; 2000MO-US16658.
PF
XX
XX 29-JUN-1999; 99US-0342364.
PR
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX
XX KIRST SJ, Holtzman DA, Fraser CC, Sharp JD, Barnes TS;
PI WPI: 2001-061966/07.
XX P-PSDB: AAB61238.
DR
XX
XX Isolated human proteins are used for diagnosis, treatment and
PT prevention of cancers, inflammatory disorders, cardiac disorders e.g.
PT arrhythmia, and skin disorders e.g. psoriasis -
XX
XX Disclosure: Fig 3I-3R; 372pp; English.
PS
XX
CC The present sequence is given in a specification relating to isolated
CC human proteins designated INTERCEPT 217, INTERCEPT 297, TANGO 276,
CC TANGO 292, TANGO 325, TANGO 331 and TANGO 332. These proteins are useful
CC as modulating agents or as targeting agents for developing agents to
CC regulate cellular processes e.g. growth, proliferation, survival,
CC differentiation and activity of human tissues. Diseases which can be
CC diagnosed, prevented and treated by administration of these polypeptides,
CC their nucleic acids and modulators include cancers, inflammatory
CC disorders, cardiac disorders e.g. arrhythmia, and skin disorders e.g.
CC psoriasis. Nucleic acids encoding the isolated proteins can be used to
CC express the proteins in a host cell in gene therapy applications.
CC Antisense molecules or ribozymes can be used to inhibit expression
CC of the proteins in target cells. Fragments of the nucleic acid molecules
CC encoding the isolated proteins are used as hybridisation probes and as
CC polymerase chain reaction (PCR) primers.
XX
XX Sequence 3503 BP; 726 A; 917 C; 1057 G; 803 T; 0 other:
SQ
Query Match 59.2%; Score 1351; DB 22; Length 3503;
Best Local Similarity 84.7%; Pred. No. 8.2e-308;
Matches 1539; Conservative 0; Mismatches 211; Indels 6; Gaps 2;
OY 127 AGACAGCTGGCCCTGAAGCTCAGACCGGGGCGTGGCCATGCGCCACACATGGGCTGTC 186
|||||
Db 61 AGGCGCTGCTGGACTGAAGTTAGACCTGGGTGCTCCATGCGCCACACATGGGCTGTC 120
OY 187 TGCTGCTGGCAGCAAGCTGTGGGGCTGGGCACTTGGGCTGAGGTGTGGAACCTT 246
|||||
Db 121 TGCTGCTGGCAGCAAGCTGTGGGGCTGGGCACTTGGGCTGAGGTGTGGAACCTT 180
OY 247 GTGCGCGGTTAAGACAGTGTCTTGGGAGCTGGCCAGGTAAGAGCGGGTCTCCAG 306
|||||
Db 181 GTGCGCGGTAAGACAGTGTCTTGGGAGCTGGTCAAGTAAGAGCGGGTCTCCAG 240
OY 307 ACCGCACTCAGAGACTTCTGACACTGACGCTGACGAGCCACCTGGGCTTCTGTAGCT 366
|||||
Db 241 ACAGCACTCAGAGACTTCTGACACTGACGCTGACGAGCAATCTTGCCCTTTATAGTG 300
OY 367 GCGCCCGAGAGGCGCTTGTGCTTACAGATGAGAGCCCTGGAGCTGCAAGAGGCATC 426
|||||
Db 301 GCGCCCGAGAGGCGCTTGTGCTTACAGATGAGAGCCCTGGAGCTGCAAGAGGCATC 360
OY 427 TCTGAGAGGCGCCCGGTGAGAGAACATGATGTATCCAGAAAGGAGAACACAG 486
|||||
Db 361 TCTTGGAGAGCTCAGCTGAGAGAAATTAATGATCCAGAAAGGAGAACACAG 420
OY 487 ACCGAGCTTCAACTTCACTCCGCTTCTGACGCTTACAGATGCTCCACCTGTAGCTC 546
|||||
Db 421 ACCGAATGCTTCAACTTCACTCCGCTTCTGACGCTTACAGATGCTCCACCTGTAGCTC 480
OY 547 TGTGCACTTACAGCTTCCAGCCCAAGTGAACCTGACCTGACATCAAGTCTACCTTGTG 606
|||||
Db 481 TGTGCACTTACAGCTTCCAGCCCAAGTGAACCTGACCTGACATCAAGTCTACCTTGTG 540
```


OY	607	GAGCATGAGAGATTGAAGATGGGAAGGGCAGATGGCTTATGACCCACTAAAGGGCCAT	566
Db	541	GACCTCGAGAAATTTGGAGATGGGAAGGGTAAATGCCCATATGACCCACTAAAGGCTCAC	600
OY	667	GCTGGCCCTCTCTTGATGATGTGAGCTGTACTCGGCCACACTCAACAATCTCTGGGCAGC	726
Db	601	ACCGGACATCCCTTGCTGGAGCGTGAAGCTGTACTACAGCCACACTCAATAAATCTCTGGGCACA	660
OY	727	GAACCCATATATCCCTCGTAAACATGGGGGGCCCACTCCATGAAGACAGATACCTGGGCC	786
Db	661	GAGCCGGGTTATCCCTTCGATACATGGGGACCCCACTCCATCAAGACAGATACCTGGCT	720
OY	787	TTTTGGCTCAACGAACTCTACTTTGTAGGCTCTGCCATATGTACTGAGAGTGTGGGCAC	846
Db	721	TTTTGGCTGAATGAACCCCACTTTGTAGGCTCTGCCCTTTGTCCCTGAGAGTGTGGGAAC	780
OY	847	TTTCAAGGGGGAGACGACGAAGGTCTACTTCTTTCAGAGGACCGGGCACTGGATCCGAC	906
Db	781	TTTCAAGGGAGAGATGACAGAAATCTACTCTTCTTTCAGAGGACCGGGCACTGGATGATGAC	840
OY	907	TGCTATGCCGAGACAGTGGTGGCTGTGTGAGCCGCTGTCTGTCGAAGGGCATATAGGGGAC	966
Db	841	TGCTATTTCCGAGACAGTGGTGGCTGTGTGAGGAGATGTGTAAAGGTATCAATGGGGGA	900
OY	967	GCACGGACCCCTGACAGAGAAAGTGCACACGTTTCTGAAAGGGCCGGCTGGCATGTCTGCC	1022
Db	901	GCACGGACGCTGCAGAAAGAAATGAGCAGACGTTTCTGAAAGGGCTGGTGTGTGCTCAGCC	960
OY	1027	CCGAATGGAGCTCTACTCTTCAACCAAGCTGACGGGGATGCATCAACCTGACAGACACCTCC	1088
Db	961	CCTGACTGGAAAGTCTACTCTCAACCAAGCTGAAAGGGGTGCACACCTGCGGGGGCCCTT	1022
OY	1087	TGGCACAACACCACTCTTCTTGGGGTTTTCAGGACAGTGGGGTGCATGTATCCTGTGC	1144
Db	1021	TGGCACAACACCACTCTTCTCGGGGTTTTCAGGAGCGGATGGGGGATGTAGACCTGTCT	1088
OY	1147	GCAATCTGTAGTATCAGATGGGAAGATCCAGGGGGTGTGAAGGGCCCTATGAAGAG	1200
Db	1081	GCAGTTTGTAGTATCAGATGGGAAGATCCAGGCAAGTGTGAAGGGTCCCTTACAAGAG	1144
OY	1207	TACCATGAGAAACCCAGAAAGTGGGACCGCTACACTGAACCTGTATCCCAAGCCCTCGGCT	1266
Db	1141	TACAGTGAAGAAACCCAGAAAGTGGGACCGCTATCTAGAACCCGATACCAAGCCCTCGGCT	1200
OY	1267	GGCTCGTGCATTTAAACAATGGCATGGGGGCCCAAGGCTACACAGCTCCCTGGAAGCTACCC	1322
Db	1201	GGTCTGTATTAACAACATGGGACCGAGACAAATGGCTACACAGTTCCTGTGAAGTGGCG	1266
OY	1327	GACAAATATCTCAACTTGTGTCAAGAAAGAACCCGCTGATAGTAGAGACAGGTGGGGCTCGG	1388
Db	1261	GACAAACACTTCAACTTGTGTCAAGAAAGAACCCCGATAGTAGAGACAGGTGAAGCTCGG	1322
OY	1387	TGAGAGCGGCCCCCTGCTCGTGAAGAAGGACACCAATTCACCCACCTGGTGGCCGACGG	1444
Db	1321	TTGGGCCCCCCCCCTACTGTGTGAAGAAACACTAATTCACACAGTGGTGGCCGACAGG	1388
OY	1447	GTTACAGGACTTGTATGAGAGCCCACTATACAGTGGCTTATTCATTTGGCAGAGAAAGCGCTGG	1500
Db	1381	GTTCCCAAGGGCTTGTATGATGGCCCACTATACAGTGTGTTCATTTGGTATCAAGAGATGGCTGG	1444
OY	1507	CTGCTCAAGGCTGTGAGGCTGGGGCCCTGGGTTTCACTGATTTGAGAGCTGCAGCTGT	1566
Db	1441	CTGCTGAAGGCTGTGAGGCTGGGGCCCTGGGATCCACATGTGTGAGGAACCTGCAGAGT	1500
OY	1567	GACCAAGAGCCCATGAGAAAGCTGGTGTATCTCAAGACAGAAAGCTGCTCTTTGGCCGG	1622
Db	1501	GACCAAGAGCCAGTGAAGAAAGTGTGTGTCTCAAGACAGAAAGTGTCTTTTCTGGCC	1566
OY	1627	TCCGCGCTTCAGCTGTGACAGTGGCCGCTGGGCGGAGCTGATTAATATGCGCTCCGTGGA	1688
Db	1561	TCCGCGCTTCAGCTGTGACAGTGTGCTGTGGCCGAGCTGCACAAATATACGTTTCTGTGTA	1622

QY	1687	GACGTGTGCTCGCCGGGAGACCCCTATTGGCCCTGGAGCGTCAACACACGCGCTGTGTG	1746
Db	1691	GACGTGTGCTCGCCAGGAGACCCCTTACTGTGCTCGTGGATGTCAACACACGCGCTGTGTG	1680
QY	1747	G----CCGTGGGTGGCCACTTTGGATCTTTTACTGATCCAGCATGTGGATGACTCGGACACT	1803
Db	1691	GCCACACCACTGTGTGCGCTCGGGGCTCTTTGTGGTCCAACTGTGGCCAACTTGTGGACACT	1740
QY	1804	TCAGGCAATTTGCAACCTCCGTGGCAGTAAAGTACAGTACAGGACCCCACTGCCAAACATC	1863
Db	1741	TCAAAGATGTGTAAACCACTATGGCATTTA-----AAAGTCAGATCTATTTCCAAACATC	1797
QY	1864	ACGGTGTGGCGGGGACAGACCTGGTGTGCTGCCCTGCACCTCTCTCTCCACTTGGCCCTG	1923
Db	1798	ACCGTTGTGTGTCAGGACACAGACCTGGTCTCTACCTGCGACCTCTGTCCAAATTTGGCCCAT	1857
QY	1924	CCCGACCTCAACCCCG	1939
Db	1858	GCCCACTGGACCTTGTG	1873

RESULT 12
AAE79451

ID AAF29451 standard; cDNA; 2811 BP.

AC AAF29451;

DT 03-APR-2001 (first entry)

DE Human TANGO 276 cDNA.

[illegible]

KW antipsoriatric; gene therapy; cancer; inflammatory disorder;

XX
XX

XX
PN
WQ200100638-A

XX 04-JAN-2001
PD

XX 16-JUN-2000: 2000WO-TS16658A
PF

XX 29-JUN-1999: 99JIS-0342364
PR

XX
PA (MILL-) MILLENNIUM PHARM INC.

PI KIRST SJ, HOLTZMAN DA, FRASER CC, SHARP JD, BARNES TS:

AA WPI; 2001-061966/07.
DR

P-PSDB; AMB01210.
XX

PT prevention of cancers, inflammatory disorders, cardiac disorders e.g. PT isolated human proteins are used for diagnosis, treatment and

XX
f1
05/05/2014, 10:00:00 e.g. 10:00:00

[illegible]

CC human proteins designated INTERCEPT 217, INTERCEPT 297, TANGO 276,

as modulating agents or as targeting agents to

CC regulate cellular processes e.g. growth, proliferation, survival,
CC differentiation and activity of human tissues. Diseases which can be

their nucleic acids and modulators include cancers, inflammatory

CC disorders, cardiac disorders e.g. arrhythmia, and skin disorders e.g. psoriasis. Nucleic acids encoding the isolated proteins can be used to

CC express the proteins in a host cell in gene therapy applications.
CC Antisense molecules or ribozymes can be used to inhibit expression

CC encoding the isolated proteins are used as hybridisation probes and a

QY	1872	GCGGGGCAACAACCTGGTGTGCCTGTGCACCTCTCCATTCAGTTGGCCTGCCGAAATC	1931
Dd	1093	GCGGGGCAACAACCTGGTGTGCCTGTGCACCTCTCCATTCAGTTGGCCTGCCGAAATC	1152
QY	1932	CMACCCCGAAG	1942
Dd	1153	GACCTTGGGG	1163

RESULT 14
ABN85379
ID ABN85379 standard; DNA; 1024 BP.

DE Partial human NOV2, semaphorin-like protein, coding sequence.
XX
KW Human: NOV2; cytosolic; Antidant; Anti-inflammatory; Immunosuppressive;
KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
KW Antiaesthetic; Nephrotoxic; Hepatotoxic; Neuroprotective; Nootropic;
KW Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
KW Gene Therapy; NOV; cancer; heart disease; Inflammation;
KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes;
KW obesity; asthma; Iga nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; infection; stroke; muscular dystrophy; epilepsy;
KW wasting disorder; semaphorin-like protein; gene; ds.

OS	Homo sapiens.	Location/Qualifiers
XX		33..1022
FH	Key	/*tag= a
FT	CDS	/partial
FT		/product= "Partial NOV2"
FT		/note= "this sequence only encodes residues 1 to 330 of NOV2, no stop codon given"
FT		

W0200255704-A2.

18-JUL-2002.

09-JAN-2002; 2002WO-US00554.

09-JAN-2001; 2001US-260417P.

28-FEB-2001; 2001US-272338P.

18-APR-2001; 2001US-284704P.

(CURA-) CURAGEN CORP.

Padlgaru M, Ll L, Z

Padigar M, Li L, Zernsen BD, Casman SJ, Shenoy S, Sptek KA, Zhong M, Gangoli EA, Burgess CE, Patturajan M, Vernet CAM; Taylor S, Tchernev V, Miller CE, Guo X, Boldog FI, Grose WM, Alsobrook JP, Gerlach V, Edinger S, Rothenberg ME, Elleman K; MacDougall J, Malyanar U, Millet I, Peyman J, Smithson G, Gunther E, Stone DJ;

WPI; 2002-590674/63.
P-PSDB; ABB98402.

NOVX polypeptides and encoding polynucleotides, useful for preventing or treating NOVX-associated disorders e.g. cancer, inflammation, or Alzheimer's disease, and in chromosome mapping, tissue typing or pharmacogenomics -

CC The present sequence is a partial coding sequence for a NOV protein. The
CC NOV proteins and coding sequences are useful for treating or preventing

CC NOV-associated disorders or in the manufacture of a medicament for
CC treating the disorders, such as cancer, heart disease, inflammation,
CC autoimmune disorders, allergies, blood disorders, AIDS, diabetes,
CC obesity, asthma, IgA nephropathy, cirrhosis, arthritis, Alzheimer's
CC disease, infections (e.g. bacterial, viral, parasitic), stroke, muscular
CC dystrophy, epilepsy, and other wasting disorders associated with chronic
CC diseases. NOV2 is a semaphorin like protein.

5Q Sequence 1024 BP; 206 A; 296 C; 320 G; 202 T; 0 other;

Query Match	40.6%;	Score 928.4;	DB 24;	Length 1024;
Best Local Similarity	92.3%;	Pred. No. 1.5e-208;		
Matches 1023; Conservative	0;	Mismatches 1;	Indels 84;	Gaps 1;

OY	134	CTGGGCTCAACCTCAGACGCGGGGGGTGCGCATG6GCCACACACTGAGGCTGTGCTGGCTG	193
Db	1	CTGGGCTCAACCTCAGACGCGGGGGGTGCGCATG6GCCACACACTGAGGCTGTGCTGGCTG	60
OY	194	TGGCAGCAGAGCTGTGGGGGCTGGGCATTTGGGGCTGAGGTGTGGTGAACCTTGTGCCG	253
Db	61	TGGCAGCAGAGCTGTGGGGGCTGGGCATTTGGGGGCTGAGGTGTGGTGAACCTTGTGCCG	120
OY	254	GTAAGACAGTGTCTTG6GGGAGCTGGGCGACGGGTGTACGGGGGTTCGCCAGACGGGCA	313
Db	121	GTAAGACAGTGTCTTG6GGGAGCTGGGCGACGGGTGTACGGGGGTTCGCCAGACGGGCA	180
OY	314	TCACAGACTTCTGTGACACTGACGCTGACGGGAGGCCACTTGGGCTTCTGTACGTGGGCGCC	373
Db	181	TCACAGACTTCTGTGACACTGACGCTGACGGGAGGCCACTTGGGCTTCTGTACGTGGGCGCC	240
OY	374	GAGAGGGCCCTGTTTTCCTTCAGCANTGAGAGCCCTGGGACTTCGAAGGACGACATCTCTGG	433
Db	241	GAGAGGGCCCTGTTTTCCTTCAGCANTGAGAGCCCTGGGACTTCGAAGGACGACATCTCTGG	300
OY	434	AGGCCCCCGTGGAGAGAAAGACTGAGTGTATCCAGAAAGGAGAAACAACAGACCGAGT	493
Db	301	AGGCCCCCGTGGAGAGAAAGACTGAGTGTATCCAGAAAGGAGAAACAACAGACCGAGT	360
OY	494	GCTTCACTTCAATCCGCTTCTCGACAGCCCTACAAATGCTCCACCTGTACGTCTGTGGCA	553
Db	361	GCTTCACTTCAATCCGCTTCTCGACAGCCCTACAAATGCTCCACCTGTACGTCTGTGGCA	420
OY	554	CCTACGCCCTCCAGCCCAAGAGCAGCTACGTACGATCAATGCTCAACTTGTGGAGCANTG	613
Db	421	CCTACGCCCTCCAGCCCAAGAGCAGCTACGTACGATCAACTTGTGGAGCANTG	480
OY	614	GAGAGTTGAAGATGGGAGAGGGCAGAGTGTCCCTATGACCACAGCTAAAGGGCCATGCTGGCC	673
Db	481	GAGAGTTGAAGATGGGAGAGGGCAGAGTGTCCCTATGACCACAGCTAAAGGGCCATGCTGGCC	540
OY	674	TTCTTGTGGATGAGTACTGTACTGTGGCCACACTACAACTTCTGTGGGACGGAAACCA	733
Db	541	TTCTTGTGGATGAGTACTGTACTGTGGCCACACTACAACTTCTGTGGGACGGAAACCA	600
OY	734	TTATTCCTGCAACATGAGGGGCCCCACACACTCCATGATGAAGACAGATGACCTTGTGGC	793
Db	601	TTATTCCTGCAACATGAGGGGCCCCACACACTCCATGATGAAGACAGATGACCTTGTGGC	660
OY	794	TCACGAACTTCACCTTGTAGGCTCTGCCATGTACTGAGAGTGGGGCAGCTTACGG	853
Db	661	TCACG-----	666
OY	854	GGGAGCAGACAAAGTCTACTTCTTTCAGGGAGCGGGGAGTGGAGTCCGACTGATG	913
Db	667	-----GGGAGCGGGCAGTGGAGTCCGACTCTATG	696
OY	914	CCGAGCAGGTGAGTGCCTGTGGCCCGTGTCTGCAAGGCGCATATG6GGGGCGACCGGA	973
Db	697	CCGAGCAGGTGAGTGCCTGTGGCCCGTGTCTGCAAGGCGCATATG6GGGGCGACCGGA	756
OY	974	CCCTGCAAGGAATGAGCACACGTTCTGTGAAGGGCGGCTGGCATGCTCTGCCCGGAAC	1033
Db	757	CCCTGCAAGGAATGAGCACACGTTCTGTGAAGGGCGGCTGGCATGCTCTGCCCGGAAC	816

```

OY 1034 GGCACCTCTACTTCAACGAGCTGCGAGGGGATGCACACCCCTGAGGACACCTCTGGCACA 1093
DB 817 GGCACCTCTACTTCAACGAGCTGCGAGGGGATGCACACCCCTGAGGACACCTCTGGCACA 876
OY 1094 ACACACCTCTCTGGGTTTTTTTCAAGCAGTGGGGTGAGCATGTACTGTGGGCATCT 1153
DB 877 ACACACCTCTCTGGGTTTTTTTCAAGCAGTGGGGTGAGCATGTACTGTGGGCATCT 936
OY 1154 GTGACTACAGTGGAGAGATCCAGCGGCTGTGGAGGGCCCTATAGAGATACATG 1213
DB 937 GTGACTACAGTGGAGAGATCCAGCGGCTGTGGAGGGCCCTATAGAGATACATG 996
OY 1214 AGGAAGCCGAGAGTGGAGCGCTACAC 1241
DB 997 AGGAAGCCGAGAGTGGAGCGCTACAC 1024

RESULT 15
AAF30193
ID AAF30193 standard; cDNA; 967 BP.
XX
XX AAF30193;
AC
XX
XX 30-APR-2001 (first entry)
XX
XX Clono 14998905.0.65 encoding SECP6.
DE
XX
XX SECP6; secreted protein; human; diagnosis; gene therapy; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH CDS 166..904
FT /*tag= a
FT sig_peptide 166..225
FT /*tag= b
FT mat_peptide 226..901
FT /*tag= c
XX
XX WO200105971-A2.
XX
XX 25-JAN-2001.
XX
XX 20-JUL-2000; 2000MO-US19890.
XX
XX 20-JUL-1999; 99US-0144722.
XX 29-NOV-1999; 99US-0167785.
XX 19-JUL-2000; 2000US-0619252.
XX
XX (CURA-) CUBAGEN CORP.
XX
XX Shinketsu RA, Fernandes E;
XX
XX MPI: 2001-091973/10.
XX P-PSDB: AAB20160.
XX
XX Now polypeptide designated SECP, its encoding nucleic acid and its
XX immunospecific antibody, useful for diagnosing, preventing and treating
XX SECP-associated disorders such as cancer -
XX
XX Claim 8; Fig 6; 124pp; English.
XX
XX The present sequence is that of a portion of clone 14998905.0.65
XX encoding novel human protein SECP6 (see AAB20160), which is
XX predicted to localise in the microbody (peroxisome). The clone was
XX obtained from lymphoid tissue. SECP6 shows identity to murine
XX hemaphysin-4c, a type I membrane protein widely expressed in the
XX nervous system during development. The invention provides 9 novel
XX SECP secreted proteins (see AAB20155-63), nucleic acids encoding
XX them (see AAF30188-96), antibodies, mutants or fragments. These
XX can be used to detect, treat or prevent an SECP-associated disorder,
XX to screen for predisposition to such a disorder, and to identify an

```

CC agent that modulates the expression or activity of SECP.

XX Sequence 967 BP; 191 A; 274 C; 288 G; 214 T; 0 other;

Query Match 36.8%; Score 841.4; DB 22; Length 967;

Best Local Similarity 99.9%; Pred. No. 4,4e-186;

Matches 842; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 CGGCCCTTCTACACATCTCTGCCCCCTGATGTGGAACGGGGTTTGGGTTTGCAGGGCT 60
DB 1 CGGCCCTTCTACACATCTCTGCCCCCTGATGTGGAACGGGGTTTGGGTTTGCAGGGCT 60
OY 61 ATTGTCTGCGCTGGGGAAGGGAGACAGCGCGGACCGGACCTCGCTGCGACCGCGCC 120
DB 61 ATTGTCTGCGCTGGGGAAGGGAGACAGCGCGGACCGGACCTCGCTGCGACCGCGCC 120
OY 121 ACCAGCAGAGACAGTGGGCTGAGAGCTGAGCGGGGGTGGGCGCATGGCCCACTGG 180
DB 121 ACCAGCAGAGACAGTGGGCTGAGAGCTGAGCGGGGGTGGGCGCATGGCCCACTGG 180
OY 181 GCTGTCTGCTGCTGCTGGCAGCAGAGCTGTGGGCTGTGGCATTTGGGGCTGAGTGTGTG 240
DB 181 GCTGTCTGCTGCTGCTGGCAGCAGAGCTGTGGGCTGTGGCATTTGGGGCTGAGTGTGTG 240
OY 241 AACCTTTGCGCGCTGAGACAGTGTCTTGTGGGAGCTGGCCAGCGTACAGCGGGTTC 300
DB 241 AACCTTTGCGCGCTGAGACAGTGTCTTGTGGGAGCTGGCCAGCGTACAGCGGGTTC 300
OY 301 TCCGAGACCGGATCCAGAGACTTCTGACACAGTGTGAGCGGAGCCCACTGGGCTTTC 360
DB 301 TCCGAGACCGGATCCAGAGACTTCTGACACAGTGTGAGCGGAGCCCACTGGGCTTTC 360
OY 361 TACGTGGGCGCCGAGAGAGCCCTGTTTGGCTTCACATGAGAGGCCCTGGAGCTGCAAGA 420
DB 361 TACGTGGGCGCCGAGAGAGCCCTGTTTGGCTTCACATGAGAGGCCCTGGAGCTGCAAGA 420
OY 421 GCGATCTCTGGGAGAGCCCGCTGGAAGAAGAGATGAGTGTATTCAGAAAGGAAGAAC 480
DB 421 GCGATCTCTGGGAGAGCCCGCTGGAAGAAGAGATGAGTGTATTCAGAAAGGAAGAAC 480
OY 481 AACGAGACCGAGTCTCAACTTCATTCGCTTCGTGACAGCCCTTACAAATGCTCCACCTG 540
DB 481 AACGAGACCGAGTCTCAACTTCATTCGCTTCGTGACAGCCCTTACAAATGCTCCACCTG 540
OY 541 TACGTCTGTGGACCTTACGCTTCGACAGCCCAAGTGCACCTAGTCAACATGCTCACCTTC 600
DB 541 TACGTCTGTGGACCTTACGCTTCGACAGCCCAAGTGCACCTAGTCAACATGCTCACCTTC 600
OY 601 ACTTTGGAGCATGAGAGATTGAAAGATGGGAAGGCAAGTGTCCCTATGAGCCACCTAAG 660
DB 601 ACTTTGGAGCATGAGAGATTGAAAGATGGGAAGGCAAGTGTCCCTATGAGCCACCTAAG 660
OY 661 GGCATGCTGGGCTTCTGTGTGATGTGAGTGTGCTGTGAGCCACATCAACAACTTCCTG 720
DB 661 GGCATGCTGGGCTTCTGTGTGATGTGAGTGTGCTGTGAGCCACATCAACAACTTCCTG 720
OY 721 GGCAGGAACCCATTATCTGCTGTAACATGAGGGGCCCCACACTCCATGAAGACAGATPAC 780
DB 721 GGCAGGAACCCATTATCTGCTGTAACATGAGGGGCCCCACACTCCATGAAGACAGATPAC 780
OY 781 CTGGCCTTTTGGCTCAACGACCTCACTTTGTAAGGCTCTGCTATGATCTGAGAGTGTG 840
DB 781 CTGGCCTTTTGGCTCAACGACCTCACTTTGTAAGGCTCTGCTATGATCTGAGAGTGTG 840
OY 841 GGC 843
DB 841 GGC 843

```

Search completed: August 19, 2003, 13:25:51
Job time : 444 secs

QY 1026 CCCGAACTGGAGCTCTACTTTCACACAGCTGAGCGGATGCAACACCTGCGAGACACTC 1085
 DB 920 ACCCGAGACTCTCCATTTCTACTTTCACAGCTGCTGAGCTGACAGGGCGTGGTACGCT 979
 QY 1086 CTGGCAACAACACACTCTTTGGGGTTTTCAGACACAGTGGGGTGAACATGACTGTC 1145
 DB 980 CGGGGGCCGGCCGCTGCTGCTGCGGCTTTTTCACGCGCCACCAACAGCATCCCTGGCTC 1039
 QY 1146 GCGCATCTGTAGTACAGTGTGAGAGATCCAGCGGGTGTGAGGGCCCTATTAAGA 1205
 DB 1040 GCGCTGTGCGGCTTGAACCTACACAGTGGAGCTGTGTTGAAGGCGGCTTCCAGA 1099
 QY 1206 GTACCATGAGAGAGCCCAAGAACTGGGAGCGCTACAC---TGACCTGTACACAGCCCTG 1262
 DB 1100 GAGAGAGTCCCGGAGTCCATCTGACGCGCGGAGATGATGAGCTGCTGACACCCG 1159
 QY 1263 GCGTGGCTGTGCTATTAACAACTGCGATGCGCGCCAGCGCTACACAGCTCCCTGAGCT 1322
 DB 1160 GCGCGGCTGCTGCGAGCGCCCGGAGTGAATGATGCTTCCAGCGCT-----T 1210
 QY 1333 ACCCGAACAACACTCTCAACTTCTGTCAGAGAGCCCGCTGATGAGAGAGAGTGGGCG 1382
 DB 1211 GCGGATGATGACATCTCAACTTGTCAAGACCACTCTGTGATGAGAGAGGCGCTCCCTC 1270
 QY 1383 TCGGTGAGAGCCCGCCCTGCTGCTGAGAGA 1411
 DB 1271 GCTGGCCATGCGCCCTGATCTGCGGA 1299

RESULT 4 US-09-077-940A-1

; Sequence 1, Application US/09077940A
 ; Patent No. 6576441
 ; GENERAL INFORMATION:
 ; APPLICANT: KIMURA, Toru et al.
 ; TITLE OF INVENTION: NOVEL SEMAPHORIN 2 AND GENE ENCODING THE SAME
 ; FILE REFERENCE: 0020-4426P
 ; CURRENT APPLICATION NUMBER: US/09/077,940A
 ; CURRENT FILING DATE: 1998-06-05
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 3692
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; NAME/KEY: 5'UTR
 ; LOCATION: (1)..(18)
 ; OTHER INFORMATION:
 ; NAME/KEY: CDS
 ; LOCATION: (19)..(2682)
 ; OTHER INFORMATION:
 ; NAME/KEY: 3'UTR
 ; LOCATION: (2683)..(3653)
 ; OTHER INFORMATION:
 ; NAME/KEY: polyA_site
 ; LOCATION: (3654)..(3692)
 ; OTHER INFORMATION:
 ; US-09-077-940A-1

Query Match 5.3%; Score 122; DB 4; Length 3692;

Best Local Similarity 51.0%; Pred. No. 6,76-20; Indels 51; Gaps 6;

Matches 479; Conservative 0; Mismatches 410;

QY 478 AACACACAGACAGAGTCTCAACTTCCAGCTTCTGACAGCCCTACAAATGCTCCAC 537
 DB 382 AAGCAAGAGGAGTGTGCGAACTTGTCAAGGTGCTCTGCTTGTGAGCAATGCAAG 441
 QY 538 CTGACGCTGTGGACCTTACAGCCCTTCCAGCCCAAGTGCACCTAGCTAGCAATGCTCAC 597
 DB 442 CTCTGTGTGGGCTCCAAATGCAATGCAATCCCATCTGTGCAATTAAGATGAGACACA 501
 QY 598 TTCACTTTTGAGCATGAGAGATTGAAGATGGGAAGGCAAGTGTCCCTATGACCCAGCT 657

DB 502 CT---GCAGCTTCTTGGAGAACAAATCAGTGTGTAAGCCCGCTGCCCCCTACAGCCCAAG 558
 QY 658 AAGGCCATGCTGGGCTCTTGTGATGTGAGCTGTACTGGGCGACACTCAAAACCTC 717
 DB 559 CATGCCAATGTGCGCTCTTCTCAGATGAGATGCTTCTACAGCCACAGTAATCTGACT 618
 QY 718 CTGGCAGGAACCCATATCTGCTTAACATGAGGGGCCACCACTGCATGAAGACAG-- 775
 DB 619 CTAGCCATTCAGAGCTGTATCTACCGTATGCGCTGGGAGCCGGCCCACTGGGACAGAG 678
 QY 776 -AGTACCTGCGCTTGTGGCTCAACAGACCTCACTTTGAGGCTGTGCTATGTACCTGAG 834
 DB 679 AAGCATGACTCCAAAGTGTGTTAAAGAGCCATCTTGTGACATGCGT----- 725
 QY 835 AGTGTGGGAGCTTACGGGGAGACAGCAAGGTCTACTTCTTCAAGAGGAGGAG 894
 DB 726 -----GAGTGGGAAAGCCACGCTCACTTCTTCCGGAGATGCGC 768
 QY 895 GTGAGTCCGACTGCTATGCCAGCAGGTGTGTGCTGTGGCCGCTGTGCAAGGCG 954
 DB 769 ATGAGATTAACTATCTGGAAGAGTGGTGTCCGCTGTGGCCGCTGTATGCAAGAT 828
 QY 955 CATATGGGGG---CGCAGGACCTGCAAGAGAGTGAACCACTTCTGAAGCGCGG 1011
 DB 839 GATGTGGGCGGCTCCCAAGGAGGTGCTGGAAGAGAGTGAATCTTCTTCAAGGCGCG 888
 QY 1012 CTGGCATGCTGTGCCCCGAACTGGCAGCTCTACTTCAACACACTGTCAGGCGATGACAG 1071
 DB 889 CTCMACTGCTCGTGTGCTGAGGAGCTCACACTTCTCAATGATCTGAGGCTGTGACT 948
 QY 1072 CTGCAAGAACACTCTTGGCAACACCAACTTCTTGGGTTTTCAGACAGTGGGCT 1131
 DB 949 GGTGTGTGAGCCTTGGGCGCGCTCCAGTATCTTGTGCTTCTCAACTCTAGCAAC 1008
 QY 1132 GACATGTACTCTCGGCCATCTGTGATGACAGCTTGAAGAGATCTCAGCGGTGTTGAG 1191
 DB 1009 AGCATCCCTGCTGCTGCTGTGCTCTTTCACATGAACCAAGTGTGCTGTGTGAA 1068
 QY 1192 GGGCCCTAATMAGAGTACCATGAGAGAGCCCAAGTGGAGCCGTACACT---GACCT 1248
 DB 1069 GCGCGCTTCCGGAGACAGAGTCACTGATCAATCTGAGACCCAGCTGTGAGACCA 1128
 QY 1249 GTACCCAGCCCTGCGCTGTGCTGTGATTAACAACCTGAGCGCCGCTGACAC 1308
 DB 1129 GTACCCAGCGCCAGCGCGGTGTGTGACAGCGCCGATATGACG-----TACAA 1179
 QY 1309 AGCTCCGTGAGCTACCGGAGCAACATCTCAACTCTGCAAGAACACCCGCTGATGAG 1368
 DB 1180 GCATCAATGCGCTTCTGACGAGATCTCAACTTGTAAAGACCCACCCACTGATGAG 1239
 QY 1369 GAGCAGTGGGCGCTGTGTGAGAGCCGCCCTGCTGTGA 1408
 DB 1240 GAGGCGTGTCCCTCCCTGAGCAACTGCGCTTGTGATTTGA 1279

RESULT 5 US-09-308-179B-2

Sequence 2, Application US/09308179B

; Patent No. 6436669
 ; GENERAL INFORMATION:
 ; APPLICANT: INAGAKI, Shinobu
 ; APPLICANT: FURUYAMA, Tatsu
 ; TITLE OF INVENTION: NOVEL SEMAPHORIN GENES (I)
 ; FILE REFERENCE: 0020-4562P
 ; CURRENT APPLICATION NUMBER: US/09/308,179B
 ; CURRENT FILING DATE: 1999-05-14
 ; PRIOR APPLICATION NUMBER: PCT/JP97/04111
 ; PRIOR FILING DATE: 1997-11-12
 ; PRIOR APPLICATION NUMBER: JAPAN 321068/1996
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.1

```

; SEQ ID NO 2
; LENGTH: 2898
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2898)
; OTHER INFORMATION: Strandedness: Double-stranded
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2898)
; OTHER INFORMATION: any n - a, c, g, t, unknown, or other
US-09-308-1795-2

```

```

Query Match      5.0%; Score 115; DB 4; Length 2898;
Best Local Similarity 48.6%; Pred. No. 3e-18;
Matches 560; Conservative 0; Mismatches 545; Indels 48; Gaps 7;

```

```

QY 310 GGCATCCAGACATTCCTGACACGCTGACGAGCCACCTGGCTTGTAGCTGGC 369
DB 529 GGAATTCCTGATCTCCATACAAATGCTGCTGATGATCAAGAACGGCTTGTGGGA 588
QY 370 GCCCGAGAGGCCCTGTTGCTTCCAGCATGGA---GGCCCTGGAGCTGCAAGAGCGATC 426
DB 589 GGCAGAGACCTTGTCTATTCCTGAACTTGAAGAGAGCTGACGCTGACAGAGATA 648
QY 427 TCCTGGAGAGCCCGCTGAGAGAAAGACTGATGTATCCAGAAAGGAGAAACACAC 486
DB 649 TACTGGCCGAGCAGACAGATGAAGATGAAGATGATGAAGG---AAAAGACGA 705
QY 487 ACCGAGTCTTCAACTTCATCCGCTTCTGACGCTTCAATGCTTCCACCTGTACGTC 546
DB 706 AATGAGTGTCCAAATATATCCGGGTTTGCATCATAACAAGACACACTTCTGACC 765
QY 547 TGTGGACACCAAGCCTTCCAGCCCAAGTGCACCTACCTCAACATG-----591
DB 766 TGTGCTAGAGACCTTGTGATCCACACTGTGCTTCCAGAGTGGGCGACCATTCAGAG 825
QY 592 ---CTGACCTTCACTTGGAGACATGAGAGATTGAAGATGGAGGCGACAGTGTCCCTAT 648
DB 826 GAACCCCTGTTTACCTGAGAGTGCACACAGATCTGAGAGAGAGAGGCGACATGCTTTT 885
QY 649 GACCCAGCTAAGAGCCATCTGCGCTTCTGTGAGTGTGAGCTGACTGACCACTC 708
DB 886 GACCCCAACTCTCTCTTGTGTCCACCTAGTGGGAATGAGCTGTGCTGACTGTAC 945
QY 709 AACAACTCTCTGGGACAGGAAACCATATATCTGCTGAATGAGGCGCCACACTCATG 768
DB 946 AGTGACTAATTTGGGGCAGAGACTGGCGATCTTCCGAGCATGGGAGATTAGGCCATAT 1005
QY 769 AAGACAGATGACTGCGCTTTTGGCTC---AAGAACCTCACTTTGTAAGCTGTGCTAT 825
DB 1006 CGCAGCTAGCATAGCATGAGCGGCTCTCTGAAGAACCAAAATTTGTAAGTCTATATG 1065
QY 826 GTACTGAGAGTGTGGGCAAGCTTCAAGGGGAGACAGCAAGGTCTACTTCTTCTGAGG 885
DB 1066 ATTCTGATTAACGAAGAC-----CGAGATGACAAACAAATGTACTTTTCTTACT 1116
QY 886 GAGCGGCACTGAGAGTCCGACTGCTATGCCGAGCAGAGTGTGGCTCTGTGGCCGCTGC 945
DB 1117 GAAGAAAGCGCTGGAGGGGAGAAACAAACGACACAGCTTACACCCGAGTGGGGCGCTG 1176
QY 946 TGCAGAGGAGATATGAGGGGCGACAGGACCTGACAGAGAGATGAGACAGTTCCTAG 1005
DB 1177 TCGGTGAATGATGATGAGAGAGACAGAGATCTGTGTGAACAGATGAGACATTTCTTAA 1236
QY 1006 GCGCGGCTGAGAGCTTGTGCCGGAACCTG-----CAGCTTACTTCAACACAGCTG 1056
DB 1237 GCGGCGCTGCTGCTAGTGGCGGGAATGATGATGATGATGATGATGATGATGATGAT 1296
QY 1057 CAGGCGATGACACACCTGACAGACACTCTGAGACAAACACACTTCTTGGGGTTT 1116
DB 1297 GAGGATGTGTTTACTGTGCGGACAGAGATCTTAAGATCCAGTGAATATTTGGACTGTT 1356

```

```

QY 1117 CAAGCAGATGGGGTGACATGTACTCTGCGCCATCTGTGAGTACAGATTGGAAGATC 1176
DB 1357 AATATACAGACATATATTTAGAGGCGACGCTGATGTGTATCATCATGTCAAGATC 1416
QY 1177 CAGCGGCTGTTTGAAGGCGCCCTATTAAGAGTACCATGAGAGAGCCAGAAAGTGGACCG 1236
DB 1417 CGGGAAGCCCTTATATGCGCCATATGCTATTAAGAAAGGCGCCCTATATACCATGTGCTA 1476
QY 1237 TACACTGACCTTATACCCAGCCCTGCGGCTGCTGCTATTAACATCTGATCGGCCG 1296
DB 1477 TATGAGAGAAAGTCCCTTACCCAGGCTGTGCTTCTGCGCAGCAAGATTAACGAGGCG 1536
QY 1297 CAGGCTACACAGACTCCCTGAGACTACCCAGACATCTCAACTTGTCTAAGAACAC 1356
DB 1537 AA---GATATGAGAACCAAGATTAACCCGATGAGCCCATCGGTTCCGAGAGATCAT 1593
QY 1357 CCGCTGATGAGAGACAGTGGGGCTCGGTGAGAGCCGCCCTGCTGTGAAGAGGCG 1416
DB 1594 CCTCTAATGTATGACGCCCATTAACCTGTTCATTAATAACCAATACTGTATTAACAGAT 1653
QY 1417 ACCAAGCTTACCC 1429
DB 1654 GGAATATCAAC 1666

```

```

RESULT 6
US-08-121-713D-53
; Sequence 53, Application US/08121713D
; Patent No. 5639856
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Mathes, David R.
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713D
; FILING DATE: 13-SEP-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; NAME/KEY: CDS
; LOCATION: 16...2331
; US-08-121-713D-53

```

Query Match 4.9%; Score 112.4; DB 1; Length 2601;
 Best Local Similarity 51.4%; Pred. No. 1.2e-17;
 Matches 356; Conservative 0; Mismatches 306; Indels 30; Gaps 3;

```

QY 358 CTGTACGTGGGGGCGCCGAGAGGCGCTGTTTGGCTTCAGCATGAGAGCCCTGGAGCTGCAG 417
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 220 CTGTATGTTGGAGCAAGAGATCAGATATTTTCTGACCTGGTAAATATCAAGATTTT 279
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 418 GGAGCATCTCTGAGAGGCGCCCGTGGAGAGAGAGCTGATATCCAGAAAGGAG 477
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 280 CAAAGATTTGTGTGGCAGATATCTTACACCAAGAGATGATGCAAGGGCTGGAGAA 339
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 478 AACAAACAGACCGAGTGTCTTCACTTCAGCTGCTTCGACAGCCCTACAGTCCCTCC 537
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 340 GACATCTCGAAGAGATGTGCTAAATTCATCAAGTACTTAAAGCATATATACAGACTAC 399
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 538 CTGTACGTCTGTGGAGACCTACGCTTCAGAGCCCAAGTGCACCTAGTCAACATGCTC--- 594
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 400 TTGTACGCTGTGGAGAGGCGGCTTTTCATCCAAATTTGACCTACATTTGAAATTTGGACAT 459
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 595 -----ACCTTCACCTTTGGAGCATGGAGATTTGAAGATTTGGAAGGCAAG 639
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 460 CATCTGAGAGCAATATTTTAAAGTGGAGAACATCAATTTTGAAGAGGCGCTGGAGAG 519
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 640 TGTCCCTATGACCCAGCTAAGGGCCATGCTGCTCTTGTGGATGTGAGCTGTACTCG 699
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 520 AGTCCATATGACCCCTAAGCTGTGACAGCATCCCTTTAAATGAGATTAATTAATACACTCT 579
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 700 GCCACATCAACAACTTCTGGGACAGGAAACCATTAATTCCTGTAACATGGGGCCACAG 759
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 580 GGAATCTGACGCTGATTTTATGGGGGAGACCTTGTCTATCTTCCGAACCTTGGGACACCAG 639
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 760 CACTCCATGAAGACAGAGTACCTGG---CCTTTTGGCTCAAGCAACCTACCTTTGTAGGC 816
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 640 CACCAATACAGGACAGAGAGATGATTCAGAGTGGCTCAATGATCCAAAGTCAATTAAGT 699
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 817 TCTGCCCTATGACCTGAGAGTGTGGGACGCTTCACAGGGGGAGACAGCAAGGCTTACTTC 876
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 700 GCCACCTCATCTCAGAGAGTGA-----CAATCTCGAAGATGACAAACTATATCTTT 750
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 877 TTCTTCAGAGGCGGAGTGGAGTCCGACTGTATGCCAGACAGTGTGGCTGTGTG 936
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 751 TTCTTCCTGTAAGATGCAATAGATGAGAAACACTCTGGAAGAGCTACTACGCTAGAAATA 810
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 937 GCCCGCTGTCTGCAAGGGCGATATGGGGGCGGACGAGCCCTGACAGAGAAATGACACAG 996
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 811 GGTGAGATATGCAAGATGACTTTTGGAGGGACAGAGAGTGTGTAATTAATGAGACAAACA 870
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 997 TTCTTCAGAGGCGGCTGGCATGCTGTGCCCC 1028
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 871 TTCTTCAGAGGCTGTGTGATTTGTCTCAGTGGC 902
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 7

US-08-835-268-53
 ; Sequence 53, Application US/08835268
 ; Patent No. 5807826
 ; GENERAL INFORMATION:
 ; APPLICANT: Goodman, Corey S.
 ; APPLICANT: Kolodkin, Alex L.
 ; APPLICANT: Matthes, David R.
 ; APPLICANT: Bentley, David R.
 ; APPLICANT: O'Connor, Timothy
 ; TITLE OF INVENTION: The Semaphorin Gene Family
 ; NUMBER OF SEQUENCES: 100
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 ; STREET: 268 Bush Street, Suite 3200
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/835,268
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/121,713
 FILING DATE: 13-SEP-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Osman, Richard A.
 REGISTRATION NUMBER: 36,627
 REFERENCE/DOCKET NUMBER: B94-002-1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415)343-4341
 TELEFAX: (415) 343-4342
 TELEX:
 INFORMATION FOR SEQ ID NO: 53:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2601 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 16..2331
 US-08-835-268-53

Query Match 4.9%; Score 112.4; DB 1; Length 2601;
 Best Local Similarity 51.4%; Pred. No. 1.2e-17;
 Matches 356; Conservative 0; Mismatches 306; Indels 30; Gaps 3;

```

QY 358 CTGTACGTGGGGGCGCCGAGAGGCGCTGTTTGGCTTCAGCATGAGAGCCCTGGAGCTGCAG 417
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 220 CTGTATGTTGGAGCAAGAGATCAGATATTTTCTGACCTGGTAAATATCAAGATTTT 279
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 418 GGAGCATCTCTGAGAGGCGCCCGTGGAGAGAGCTGATATCCAGAAAGGAG 477
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 280 CAAAGATTTGTGTGGCAGATATCTTACACCAAGAGATGATGCAAGGGCTGGAGAA 339
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 478 AACAAACAGACCGAGTGTCTTCACTTCAGCTGCTTCGACAGCCCTACAGTCCCTCC 537
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 340 GACATCTCGAAGAGATGTGCTAAATTTTCAAGAGTACTTAAAGCAATTAATGAGACTAC 399
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 538 CTGTACGTCTGTGGAGACCTACGCTTCACAGCCCAAGTGCACCTACGTCACATGCTC--- 594
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 400 TTGTACGCTGTGGAGAGGCGGCTTTTCATCCAAATTTGACCTTACATTTGAAATTTGACAT 459
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 595 -----ACCTTCACCTTTGGAGCATGGAGATTTGAAGATTTGGAAGGCAAG 639
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 460 CATCTGAGAGCAATATTTTAAAGTGGAGAACATCAATTTTGAAGAGGCGCTGGAGAG 519
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 640 TGTCCCTATGACCCAGCTAAGGGCCATGCTGCTCTTGTGGATGTGAGCTGTACTCG 699
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 520 AGTCCATATGACCCCTAAGCTGTGACAGCATCCCTTTAAATGAGAGGAATTAATCTCT 579
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 700 GCCACATCAACAACTTCTGGGACAGGAAACCATTAATTCCTGTAACATGGGGCCACAG 759
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 580 GGAATCTGAGCTGATTTTATGGGGGAGACTTTTGTCTATCTTCCGAACCTTGGGACACAG 639
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 760 CACTCCATGAAGACAGAGTACCTGG---CCTTTTGGCTCAAGCAACCTACCTTTGTAGGC 816
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 640 CACCAATACAGGACAGAGAGATGATTCAGAGTGGCTCAATGATCCAAAGTCAATTAAGT 699
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 817 TCTGCCCTATGACCTGAGAGTGTGGGACGCTTCACAGGGGGAGACAGCAAGGCTTACTTC 876
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 700 GCCACCTCATCTCAGAGAGTGA-----CAATCTCGAAGATGACAAACTATATCTTT 750
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 877 TTCTTCAGAGGCGGAGTGGAGTCCGACTGTATGCCAGACAGTGTGTGCTGTGTG 936
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 751 TTTCTCCGTAAGAAATGATGAGAACACTCTGGAAAAGCTACTACAGTAGATA 810
QY 937 GCCCGTGTCTGCAAGGCGATATGGGGGCGACAGCCCTGAGAGAAATGGACG 996
Db 811 GGTGAGATGATGCAAGATGACTTTGGAGGGGACAGAGTCTGTGATTAATGACACA 870
QY 997 TTCTGAAAGCGCGCTGGCATGCTGCCCC 1028
Db 871 TTCTCAAGCTGCTGTGATTTGCTCAGTGCC 902

RESULT 8

US-09-060-692-53
Sequence 53, Application US/09060692
Patent No. 5935865

GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Mathes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,692
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/121,713
FILING DATE: 13-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 2601 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 16..2331
US-09-060-692-53

Query Match 4.9% Score 112.4; DB 2; Length 2601;
Best Local Similarity 51.4%; Pred. No. 1.2e-17;

Matches 356; Conservative 0; Mismatches 306; Indels 30; Gaps 3;

QY 358 CTGTACGTGGGCGCGGAGAGCGCTGTGGCTTCAGCATGAGCGCTGAGCTGCAA 417
Db 220 CTGTATGTGGAGCAAGGATCATATTTTCATTCGACGCGTAAATATCAAGATTTT 279
QY 418 GGACGATCTCTCTGGAGCGCCCGCTGGAGAGAGACTGATTCACAGAAAGGAG 477

Db 280 CAAAAGATTGTGTGCCAGTATCTTACACCAAGAGATGAAATGCAAGTGGCTGGA 339
QY 478 AACACAGAGACGAGTCTTCAACTTATCCGCTTCTGAGCGCTCAATAGCTCCAC 537
Db 340 GACATCTGAAAGATTTGTAATTTATCATAGGTACTTAAAGATTAATATGACGAC 399
QY 538 CTGTACGTGTGGGACCTTACGCTTCCAGGCCAAGTGCACCTACGTAACATGCTC 594
Db 400 TTGTACGCTGTGGAGAGCGGGCTTTTCATTCATCAATTTGCACTCATTTGAAT 459
QY 595 -----ACCTCATTGGAGCATGGAGTTGAAGATGGAGGAG 639
Db 460 CATCTGAGCAATATTTTAACTGAGAACTCACATTTTGAAGCGCGTGGAG 519
QY 640 TGCCCTATGACCCAGTAAAGGCGCATGCGCTTCTGTGATGATGAGCTGACTCG 699
Db 520 AGTCATATGACCTTAAGCTGCTACAGCATCCCTTTTAATAGTGAATTTACTCT 579
QY 700 GCCACACTCAACAACTTCTGGGCGAGAACCCATTAATCTGCTAACATGGGCGCC 759
Db 580 GGAAGTGCAGCTGATTTATGGGCGAGACTTGTCTATCTCCGAACTTGGGCGAC 639
QY 760 CACTCCATGAAAGAGATGACTG---CCTTTGGCTCAACGAACTCACTTGTAGC 816
Db 640 CACCAATCAGACAGACAGCATGATTCAGGTGCTCATGATCCAAAGTTGATAGT 699
QY 817 TCTCCTATGATGACCTGAGAGTGTGGGCGAGCTTACAGGGGAGACAGCAAGGTCT 876
Db 700 GCCACCTCATCTAGAGAGTGA-----CAATCTGAAGATGACAAAGTATTT 750
QY 877 TTCTCAAGGAGCGGCGAGTGGAGTCCGACTATGCCGAGAGTGTGCTGCTGTG 936
Db 751 TTTCTCCGTGAAATGCAATGATGAGAAACACTCTGGAAGAGTACTACAGCTAGA 810
QY 937 GCCCGTGTCTGCAAGGCGATATGGGGGCGACAGGACCTGCGAGAGAAATGAC 996
Db 811 GGTGAGATGATGCAAGATGACTTTGGAGGGGACAGAAAGTGTGTAATTAATGACA 870
QY 997 TTCTGAAAGCGCGCTGGCATGCTGCCCC 1028
Db 871 TTCTCAAGCTGCTGTGATTTGCTCAGTGCC 902

RESULT 9

US-08-833-391-53
Sequence 53, Application US/08833391
Patent No. 6013781

GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Mathes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,391
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/121,713

Db 460 CATCTGAGAGCAATATTTTAAAGCTGAGAACTCAATTTTGAAGAGCCCTGGGAG 519
QY 640 TGTCCCTATGACCCAGCTAAGGCGCATGCTGAGCTTCTGTGGATGGTGTACTG 699
Db 520 AGTCCATATGACCCCTAAGCTGTGACAGCATCCCTTTATATGATGAGATATATCT 579
QY 700 GCCACACTCAACACTTCTGAGGACGAGAACCCATTAATCTGCTAATAGGGGCCAC 759
Db 580 GGAATGACAGCTATTTTATGAGGCGGAGCTTGTATCTTCGAACTTGGGACAC 639
QY 760 CACTCATGAGACAGATGCTG---CCTTTGGCTCAAGACCTCTTTGAGC 816
Db 640 CACCCATATGACGACAGCATGATTCAGGTGGCTCAATGATCCAAAGTTCATGCT 699
QY 817 TCTGCTATGCTAGAGTGTGGGAGCTTCAAGGGGAGGACGACGACGATCTTCT 876
Db 700 GCCCACTCATCTCAGAGAGTGA-----CAATCTGAAGATGACAAAGTATCTT 750
QY 877 TTCTCAGGAGGCGGAGTGAAGTCCGACTGCTATGCCGAGAGGTGTGCTGTG 936
Db 751 TTCTCCGTAAATGCAATGATGAGAACACTCTGGAAGGCTACTCAGGCTAGATA 810
QY 937 GCCCGTCTGCAAGGCGCATATGAGGGGCGGACGACCTCGAGAGAGTGGACGAC 996
Db 811 GGTCAATATGCAAGATGACTTTGAGGGGACAGAGTCTGTGAATTAATGACACAA 870
QY 997 TTCTGAAGGCGGCTGGCATGCTGCCCC 1028
Db 871 TTCTCAAGCTGCTGATTTGCTCAGTGCC 902

RESULT 11
PCT-US94-10151A-53

; Sequence 53, Application PC/TUS9410151A
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR HOHBACH TEST ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10151A
; FILING DATE: 13-SEP-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: EP-58750-PC/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..2331

PCT-US94-10151A-53

Query Match 4.9%; Score 112.4; DB 5; Length 2601;
Best Local Similarity 51.4%; Pred. No. 1,2e-17;
Matches 356; Conservative 0; Mismatches 306; Indels 30; Gaps 3;

QY 358 CTGTACGTGGGCGCGGAGAGGCGCTGTGTTGCTTCAGATGAGGCCCTGAGCTGCA 417
Db 220 CTGTATGTTGAGCAAGATGATCAATATTTTCAATGACCTGGTATATCAAGATTTT 279
QY 418 GGAGCGATCTCTGAGGAGGCGCCCGTGGAGAAAGACTAGTATCCAGAAAGGAG 477
Db 280 CAAAGATGTGTGGCCAGTATCTTACACCAAGAGAGATTAAGTCAAGTGGCTGGAAA 339
QY 478 AACACAGACAGAGTCTTCAACTCATCTCCGCTTCTGAGGCCCTCAATGCTCCAC 537
Db 340 GACATCTGAAAGATGTGCTATTTCATCAAGTACTTAAGCATATTAATCAGACTCAC 399
QY 538 CTGTACGTGTGGACCTACGCTTCACGCCCAAGTGCACCTACGTCACATGCTC 594
Db 400 TTGTACGCTGTGGAGGCGGCTTTTCATCAATTTGACACTACATATGAATGACAT 459
QY 595 -----ACCTTCACTTTGAGACATGAGAGTTGAAGATGGGAGGACG 639
Db 460 CATCTGAGAGCAATATTTTAAAGCTGAGAACTCAATTTTGAAGAGCCCTGGGAG 519
QY 640 TGTCCCTATGACCCAGCTAAGGCGCATGCTGAGCTTCTGTGATGGTGTGACTGCTG 699
Db 520 AGTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAATGATGAGATATATCT 579
QY 700 GCCCACTCAACACTTCTGAGGACGAGAACCCATTAATCTGCTAATATGAGGCCAC 759
Db 580 GGAAGTCACTGATTTTATGAGGGGCGAGACTTGTCTATCTCCGAACTTGGGACAC 639
QY 760 CACTCATGAGAGAGAGTACTG---CCTTTGGCTCAAGAACTCACTTGTAGC 816
Db 640 CACCAATGAGAGAGAGCATGATTCAGGCTGCATGATCCAAAGTCTATTAGT 699
QY 817 TCTGCTATGCTAGCTGAGAGTGTGGGAGCTTCAAGGGGAGGACGACAGAGTCTACT 876
Db 700 GCCCACTCATCTCAGAGAGTGA-----CAATCTGAAGATGACAAAGTATATTT 750
QY 877 TTCTCAGGAGGCGGAGTGTGAGTCCGACTGTATGCCGAGAGGTGTGCTGTGTG 936
Db 751 TTCTCCGTAAATGCAATGATGAGAGAACTCTGGAAGGCTACTCAGCTAGATA 810
QY 937 GCCCGTCTGCAAGGCGCATATGAGGGGCGGACGACCTCGAGAGAGAGTGGACGAC 996
Db 811 GGTCAATATGCAAGATGACTTTGAGGGGACAGAGTCTGTGAATTAATGACACAA 870
QY 997 TTCTGAAGGCGGCTGGCATGCTGCCCC 1028
Db 871 TTCTCAAGCTGCTGATTTGCTCAGTGCC 902

RESULT 12

US-09-254-594-5
; Sequence 5, Application US/09254594
; Patent No. 6566094
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru
; APPLICANT: KIKUCHI, Kaoru
; TITLE OF INVENTION: NOVEL SEMAPHORIN GENE: SEMAPHORIN Y
; FILE REFERENCE: 0020-4527P
; CURRENT APPLICATION NUMBER: US/09/254,594
; CURRENT FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 5
; LENGTH: 2790
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

;; CURRENT FILING DATE: 1999-05-11
;; NUMBER OF SEQ ID NOS: 13
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 1
;; LENGTH: 3195
;; TYPE: DNA
;; ORGANISM: Rattus norvegicus
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)..(50)
;; OTHER INFORMATION: Identification Method: E
;; NAME/KEY: 5'UTR
;; LOCATION: (1)..(50)
;; OTHER INFORMATION: Identification Method: E
;; NAME/KEY: misc_feature
;; LOCATION: (51)..(2837)
;; OTHER INFORMATION: CDS; Identification Method: E
;; NAME/KEY: 3'UTR
;; LOCATION: (2838)..(3195)
;; OTHER INFORMATION: Identification Method: E
US-09-254-594-1

Query Match 3.4%; Score 77.4; DB 4; Length 3195;
Best Local Similarity 46.8%; Pred. No. 3.3e-09;
Matches 314; Conservative 0; Mismatches 351; Indels 6; Gaps 2;

OY 852 GGGGAGCAGCAGCAAGTCTACTTCTTCTTCTGAGGAGCGGCAGT--GAGTCCGACTG 908
DB 737 GGAGCATGAGAGCACTGCTACTTCTTCTCCGAGAGAGTCTGTGTGAGAGACCCGG 796
OY 909 CTATGCCAGCAGAGGTGGTCTGTGTGGCCGCTCTGCAGAGGCGATATGGGGGC-- 966
DB 797 CCTGGGAGGGGTGCAATTTCCCGGGTGGCCGGGTGTAAACGTGACATGGGTGGCTC 856
OY 967 -GCAGGACCTGCGAGAGAGTGGAGCAGTCTCTGAAGGGCGCGTGGCATGCTCTGC 1025
DB 857 ACCACGGGCTTGTGATGCCACTGACATCCTTCTTAAGCTGAGGCTCAACTGCTCCGT 916
OY 1026 CCCGAACTGGGAGCTTACTTCAACCACTGAGGCGATGCAACCCCTGCAGAGACCTC 1085
DB 917 CCTGCGGACTTACTTCTTACTTGTATGCTTACAGTCTTAAGTGGCCGTGTACCT 976
OY 1086 CTGGCACAACACACTTCTTTGGGTTTTCAGACAGTGGGTGACATGTACTGTC 1145
DB 977 GCATGGCGGCTGTGCGCTTTGGGTTCTCACTACAGCAATAGCATTCCTGGGTC 1036
OY 1146 GGCAATCTGTAGTACAGTTCAGAGATCCAGCGGCTTTGAGGGCCCTATATAGGA 1205
DB 1037 TGCAGTCTGGCGCTTCTACCTAGATGACATGAGAGTGGCTTTGAGGCAATTCAGGA 1096
OY 1206 GTACCATGAGAGCCCAAGAGTGGGACCGCTACACTGACCCCTGACCCGCTGGCC 1265
DB 1097 GCAGAGGAGTGTGTGGGCGCTGAGTCCTGTGTGAGGACAAAGTCCCTCACCCAG 1156
OY 1266 TGGCTCGTGCATTAACAAGTGGATCGGCGCAGGCTACACAGCTCCCTGGAGCTACC 1325
DB 1157 GCCAGGGTCTGTGAGGTGTGGGTGAGCTGCTTATTCCTCTCTCAAGACTGCC 1216
OY 1326 CGACAACATCTCAACTGTCAGAGAACACCCGCTGATGAGAGAGAGAGTGGGCGCTCG 1385
DB 1217 TGACGATGTCTGCTCTTCAATCAAGGACACACCACTGCTGGATCCCGCTGTGCCACTGC 1276
OY 1386 GTGAGCGCCGCTGCTGCTGTAAGAGGAGCACCACCACTTCAACCCAGCTGTGGCCGACCG 1445
DB 1277 CACCCATCAACCTCTCCCTCACTGACTGACAGGCTCTACTGAACCCAGGTAGCTGTGA 1336
OY 1446 GGTACAGGACTTGTAGGAGCACCACCTATACAGTGTCTTCAATTGGCACAGAGACGGCTG 1505
DB 1337 TGGTATGGCTGGCCGCCAGAGAAATACTACAGTCTCTTCTGTGCTCCAATGATGGGAC 1396
OY 1506 GCTGCTCAAG 1516
DB 1397 AGTCTGAAG 1407

Search completed: August 19, 2003, 15:55:01
Job time : 125 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 14:58:57 ; Search time 418 Seconds
(without alignments)
12228.363 Million cell updates/sec

Title: US-10-002-050-13

Perfect score: 2284
Sequence: 1 cggcccttcacactctctg.....tctgtgtgtatgtgtgtctg 2284

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1504479 seqs, 1118970152 residues

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2280	99.8	2284	13	US-10-003-152-13 Sequence 13, Appl
2	2280	99.8	2284	14	US-10-002-050-13 Sequence 13, Appl
3	2280	99.8	2284	14	US-10-002-304-13 Sequence 11, Appl
4	1887.4	82.6	2155	13	US-10-003-152-11 Sequence 11, Appl
5	1887.4	82.6	2155	14	US-10-002-050-11 Sequence 11, Appl
6	1887.4	82.6	2155	14	US-10-002-304-11 Sequence 11, Appl
7	1879.4	82.3	2156	13	US-10-003-152-21 Sequence 21, Appl
8	1879.4	82.3	2156	14	US-10-002-050-21 Sequence 21, Appl
9	1879.4	82.3	2156	14	US-10-002-304-21 Sequence 21, Appl
10	1761	77.1	3293	14	US-10-149-819-25 Sequence 25, Appl
11	1394	61.0	8095	10	US-09-989-920-73 Sequence 73, Appl
12	1351	59.2	3503	11	US-09-759-130B-336 Sequence 336, Appl
13	1351	59.2	3503	14	US-10-189-123-66 Sequence 66, Appl
14	1193.6	52.3	2811	11	US-09-759-130B-303 Sequence 303, Appl
15	1193.6	52.3	2811	11	US-10-189-123-33 Sequence 33, Appl
16	841.4	36.8	967	13	US-10-000-512-11 Sequence 11, Appl

17	463	20.3	729	11	US-09-759-130B-304	Sequence 304, App
18	463	20.3	729	14	US-10-189-123-34	Sequence 34, Appl
19	411	18.0	3781	11	US-09-946-374-252	Sequence 252, Appl
20	411	18.0	3781	12	US-10-015-387A-252	Sequence 252, App
21	411	18.0	3781	12	US-10-006-130A-252	Sequence 1, Appl
22	411	18.0	3781	12	US-10-199-672-453	Sequence 453, App
23	411	18.0	3781	12	US-10-241-220-1	Sequence 453, App
24	411	18.0	3781	13	US-10-052-586-453	Sequence 453, App
25	411	18.0	3781	14	US-10-174-590-453	Sequence 453, App
26	411	18.0	3781	14	US-10-176-758-453	Sequence 453, App
27	411	18.0	3781	14	US-10-175-737-453	Sequence 453, App
28	411	18.0	3781	14	US-10-173-706-453	Sequence 453, App
29	411	18.0	3781	14	US-10-175-738-453	Sequence 453, App
30	411	18.0	3781	14	US-10-175-752-453	Sequence 453, App
31	411	18.0	3781	14	US-10-176-482-453	Sequence 453, App
32	411	18.0	3781	14	US-10-176-757-453	Sequence 453, App
33	411	18.0	3781	14	US-10-176-913-453	Sequence 453, App
34	411	18.0	3781	14	US-10-180-552-453	Sequence 453, App
35	411	18.0	3781	14	US-10-180-557-453	Sequence 453, App
36	411	18.0	3781	14	US-10-180-557-453	Sequence 453, App
37	411	18.0	3781	14	US-10-174-572-453	Sequence 453, App
38	411	18.0	3781	14	US-10-174-579-453	Sequence 453, App
39	411	18.0	3781	14	US-10-174-582-453	Sequence 453, App
40	411	18.0	3781	14	US-10-174-588-453	Sequence 453, App
41	411	18.0	3781	14	US-10-175-739-453	Sequence 453, App
42	411	18.0	3781	14	US-10-175-740-453	Sequence 453, App
43	411	18.0	3781	14	US-10-175-743-453	Sequence 453, App
44	411	18.0	3781	14	US-10-176-488-453	Sequence 453, App
45	411	18.0	3781	14	US-10-176-492-453	Sequence 453, App

ALIGNMENTS

RESULT 1
US-10-003-152-13
Sequence 13, Application US/10003152
Publication No. US20020151494A1
GENERAL INFORMATION:
APPLICANT: Shinketsu, Richard
APPLICANT: Fernandes, Elma
APPLICANT: Vernet, Corine
APPLICANT: Yang, MeiJia
APPLICANT: Boldog, Ferenc
TITLE OF INVENTION: No. US20020151494A1 Amino Acid Sequences for Human Semaphorin
FILE REFERENCE: 15966-554 Cura-54 CON-S12
CURRENT APPLICATION NUMBER: US/10/003,152
CURRENT FILING DATE: 2001-11-02
PRIOR APPLICATION NUMBER: 09/604,286
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/140,584
PRIOR FILING DATE: 1999-06-23
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 2284
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (166)..(1953)
NAME/KEY: variation
LOCATION: (1)..(2284)
OTHER INFORMATION: N may be any nucleotide
US-10-003-152-13

Query Match 99.8% Score 2280; DB 13; Length 2284;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGCCCTTCACACTCCTGCCCTGATGTGAGACGGGGTTGGGGTTCTGCAGGCT 60
|||||

Db 1 CGGCCCTTCTACACATCTCCCTGCTGATGTGGAGCGGGGTTTGGGGTTCTGCAAGGCT 60
Oy 61 ATTGCTGCGCTGGGGAAGGGAGAGGCCGGGAGCCCTCGCTGCGACGCGGCC 120
Db 61 ATTGCTGCGCTGGGGAAGGGAGAGGCCGGGAGCCCTCGCTGCGACGCGGCC 120
Oy 121 ACCAGCAGGAGAGCTGGGCTGAGAGCTCAGAGCCGGGGGCTGCGCTAGTGGCCCACTGG 180
Db 121 ACCAGCAGGAGAGCTGGGCTGAGAGCTCAGAGCCGGGGGCTGCGCTAGTGGCCCACTGG 180
Oy 181 GCTGCTGCTGCTGCGCAGCAAGGCTGAGGGCCCTGGGCAATTGGGGGCTGAGTGGTGG 240
Db 181 GCTGCTGCTGCTGCGCAGCAAGGCTGAGGGCCCTGGGCAATTGGGGGCTGAGTGGTGG 240
Oy 241 AACCTTGCGCGGCTAAGACAGTGTCTTGCGGAGCTGGCCAGCGTAGTACGGCGGTTG 300
Db 241 AACCTTGCGCGGCTAAGACAGTGTCTTGCGGAGCTGGCCAGCGTAGTACGGCGGTTG 300
Oy 301 TCCGAGACCGGCAATCCAGAGATTCCTGACACTGACGCTGAGAGGAGCCCACTGGGCTTCTG 360
Db 301 TCCGAGACCGGCAATCCAGAGATTCCTGACACTGACGCTGAGAGGAGCCCACTGGGCTTCTG 360
Oy 361 TAGCTGGGCGCCGAGAGAGCCCTGTTTGCCTTCAGACATGAGAGCCCTGGAGCTGCAAGA 420
Db 361 TAGCTGGGCGCCGAGAGAGCCCTGTTTGCCTTCAGACATGAGAGCCCTGGAGCTGCAAGA 420
Oy 421 GCGATCTCTCTGGAGAGCCCGCGTGAGAGAGACAGTGAAGTATCCAGAAAAGGAGAAC 480
Db 421 GCGATCTCTCTGGAGAGCCCGCGTGAGAGAGACAGTGAAGTATCCAGAAAAGGAGAAC 480
Oy 481 AACAGACCGAGTGTCAATCAATCCGCTTCGAGGCGCTCAATGCGCTCCACCTG 540
Db 481 AACAGACCGAGTGTCAATCAATCCGCTTCGAGGCGCTCAATGCGCTCCACCTG 540
Oy 541 TAGCTGTGTGACCTTACCTAGCCCTTCAGCCCAAGTACACCTAGTCAACATGCTCACTTC 600
Db 541 TAGCTGTGTGACCTTACCTAGCCCTTCAGCCCAAGTACACCTAGTCAACATGCTCACTTC 600
Oy 601 ACTTGGAGCATGAGAGATTGAAAGTGGAAAGGCAAGTGTCCCTATGACCCAGCTAAG 660
Db 601 ACTTGGAGCATGAGAGATTGAAAGTGGAAAGGCAAGTGTCCCTATGACCCAGCTAAG 660
Oy 661 GGCATGTGTGCTTCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Db 661 GGCATGTGTGCTTCTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Oy 721 GGCAGGGAACCATTTATCTGCTAACAATGGGGGCCCACTCCATGATGAAGACAGATAC 780
Db 721 GGCAGGGAACCATTTATCTGCTAACAATGGGGGCCCACTCCATGATGAAGACAGATAC 780
Oy 781 CTGGGCTTTTGGCTACAGAACTTCATTTGATGGCTTGGCTATGATCTGAGAGTGG 840
Db 781 CTGGGCTTTTGGCTACAGAACTTCATTTGATGGCTTGGCTATGATCTGAGAGTGG 840
Oy 841 GGCAGCTTACAGGGGGAGAGAGAGAGTCTACTCTTTCAGGGAGCGGGAGTGGAG 900
Db 841 GGCAGCTTACAGGGGGAGAGAGAGAGTCTACTCTTTCAGGGAGCGGGAGTGGAG 900
Oy 901 TCCGACTGCTATGCGAGCAGGTGTGCTGATGGGCCGTGTCTGCAAGGGCGCATATG 960
Db 901 TCCGACTGCTATGCGAGCAGGTGTGCTGATGGGCCGTGTCTGCAAGGGCGCATATG 960
Oy 961 GGGGGGAG 1020
Db 961 GGGGGGAG 1020
Oy 1021 TCTGCGCCGAACTGGAGAGCTTACTTCAACAGAGCTGCAAGGGGATGACACCTGAGAGAC 1080
Db 1021 TCTGCGCCGAACTGGAGAGCTTACTTCAACAGAGCTGCAAGGGGATGACACCTGAGAGAC 1080
Oy 1081 ACCTCTGCGCAACACCACTCTCTTGGGGTTTTTCAAGCAGAGTGGGATGACATGTAC 1140
Db 1081 ACCTCTGCGCAACACCACTCTCTTGGGGTTTTTCAAGCAGAGTGGGATGACATGTAC 1140

Oy 1141 CTGCGCCATCTGTGATACCAAGTTGGAAGATTCAGCGGGTGTGTTGAGGGCCCTAT 1200
Db 1141 CTGCGCCATCTGTGATACCAAGTTGGAAGATTCAGCGGGTGTGTTGAGGGCCCTAT 1200
Oy 1201 AAGGATGCATGAG 1260
Db 1201 AAGGATGCATGAG 1260
Oy 1261 CGGCTGCTGCTGATTAACAATGCGATCGGCGCACAGGCTACACCAAGCTCCCTGGAG 1320
Db 1261 CGGCTGCTGCTGATTAACAATGCGATCGGCGCACAGGCTACACCAAGCTCCCTGGAG 1320
Oy 1321 CTACCGACAAATCTCACTTGTCAAGAGACCCGCTGATGAGAGAGAGAGAGAGAGAG 1380
Db 1321 CTACCGACAAATCTCACTTGTCAAGAGACCCGCTGATGAGAGAGAGAGAGAGAGAG 1380
Oy 1381 CCTGGTGAAGCGCGCCCTGCTGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1381 CCTGGTGAAGCGCGCCCTGCTGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Oy 1441 GACCGGTTAAG 1500
Db 1441 GACCGGTTAAG 1500
Oy 1501 GCGTGGCTGCTCAG 1560
Db 1501 GCGTGGCTGCTCAG 1560
Oy 1561 CTGTTGACAG 1620
Db 1561 CTGTTGACAG 1620
Oy 1621 GCGGCTGCTCAG 1680
Db 1621 GCGGCTGCTCAG 1680
Oy 1681 TGTGCAAGCTGTGTCTTGTGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 1681 TGTGCAAGCTGTGTCTTGTGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Oy 1741 TGTGCGCGGAG 1800
Db 1741 TGTGCGCGGAG 1800
Oy 1801 ACTTCAGGCAATTTGCAACCTCGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Db 1801 ACTTCAGGCAATTTGCAACCTCGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Oy 1861 ATCAGGAG 1920
Db 1861 ATCAGGAG 1920
Oy 1921 CTGCGGAG 1980
Db 1921 CTGCGGAG 1980
Oy 1981 ANAGGAG 2040
Db 1981 ANAGGAG 2040
Oy 2041 ACGGAG 2100
Db 2041 ACGGAG 2100
Oy 2101 GCTGCTGCTCAG 2160
Db 2101 GCTGCTGCTCAG 2160
Oy 2161 CATCTTCT 2220
Db 2161 CATCTTCT 2220

[illegible][illegible]

QY	781	TTGGCCTTTTGGCCCAACGAAACCTCATTTGTAGGCGTCGGCTATGTACCTGACAGTGTG	840
Dp	781	CTGGCCTTTTGGCTTCAACGAAACCTCATTTGTAGGCTGTGCTATGTACCTGACAGTGTG	840
QY	841	GGCAGCTTCAAGGGGGGAGCAGACAGAGGTCTACTTCTTCTCAGAGGAGCGGGCAGTGGAG	900
Dp	841	GGCAGCTTCAAGGGGGGAGCAGACAGAGGTCTACTTCTTCTCAGAGGAGCGGGCAGTGGAG	900
QY	901	TCGCAGTGTCTATGCCGAGCAGAGTGTGGCTCTGTGGCCCTGTCTGTCAMAGGGCGATATG	960
Dp	901	TCGCAGTGTCTATGCCGAGCAGAGTGTGGCTCTGTGGCCCTGTCTGTCAMAGGGCGATATG	960
QY	961	GGGGGGCCAGAGACCCCTGCAGAGGAATGGACACAGTTCCTCGAAGGGCGGGCTGGCATGCG	1020
Dp	961	GGGGGGCCAGAGACCCCTGCAGAGGAATGGACACAGTTCCTCGAAGGGCGGGCTGGCATGCG	1020
QY	1021	TCCTCCCGGAAGCTGGCAGCTCTACTTCAACAGCTGAGAGGGGATGSCAACCTTCAGAGAC	1080
Dp	1021	TCCTCCCGGAAGCTGGCAGCTCTACTTCAACAGCTGAGAGGGGATGSCAACCTTCAGAGAC	1080
QY	1081	ACCTCTGTGGCACAACACACCTTCTTTGGGGTTTTTCAGACAGCAGTGGGGGTGACATGTAC	1140
Dp	1081	ACCTCTGTGGCACAACACACCTTCTTTGGGGTTTTTCAGACAGCAGTGGGGGTGACATGTAC	1140
QY	1141	CTGTGCGGCATCTGTGAGTACCACTTTGGAGAGATCCAGCGGGGTGTTTGAAGGCCCTAT	1200
Dp	1141	CTGTGCGGCATCTGTGAGTACCACTTTGGAGATCCAGCGGGGTGTTTGAAGGCCCTAT	1200
QY	1201	AAGAGATACCATGAGGAAAGCCCAAGGGGAGCGGTACACTGACCCGTATGCCAGGCCT	1260
Dp	1201	AAGAGATACCATGAGGAAAGCCCAAGGGGAGCGGTACACTGACCCGTATGCCAGGCCT	1260
QY	1261	CGGCGCTGGCTCTGATTAACAACTGGGATCGGGCCACGGCTACACAGCTTCCTGTGAG	1320
Dp	1261	CGGCGCTGGCTCTGATTAACAACTGGGATCGGGCCACGGCTACACAGCTTCCTGTGAG	1320
QY	1321	CTACCCGCAACATCCCTCAACTTGTCTCAAGAGACCCGGCTGATGAGAGAGCAGTGGGG	1380
Dp	1321	CTACCCGCAACATCCCTCAACTTGTCTCAAGAGAGCCGGCTGATGAGAGAGCAGTGGGG	1380
QY	1381	CCTCGGTGAGACCCGCCCTCTGCTGTAAGAGGGGACACAACTTCACCCACTCTGTGGCC	1440
Dp	1381	CCTCGGTGAGACCCGCCCTCTGCTGTAAGAGGGGACACAACTTCACCCACTCTGTGGCC	1440
QY	1441	GACCGGGTTACAGGACTTGATGAGACCACCTATACAGTGGTTCATTGTGCGACAGGAGAC	1500
Dp	1441	GACCGGGTTACAGGACTTGATGAGACCACCTATACAGTGGTTCATTGTGCGACAGGAGAC	1500
QY	1501	GGCTGGCTGCTCAAGGCTGTGAGCCTGGGGCCCTGGGTTCACTGATTTAGAGAGCTGCAG	1560
Dp	1501	GGCTGGCTGCTCAAGGCTGTGAGCCTGGGGCCCTGGGTTCACTGATTTAGAGAGCTGCAG	1560
QY	1561	CTGTTTACAGAGAGCCCATAGAAAGCTGGTGTATCTCAGACCAAGAGCTGCTCTT	1620
Dp	1561	CTGTTTACAGAGAGCCCATAGAAAGCTGGTGTATCTCAGACCAAGAGCTGCTCTT	1620
QY	1621	GCCGCGTCCGCTCTCAGCTGTGAGCTGCGCTGGCGGACGTCATTAAGTATGCGTCC	1680
Dp	1621	GCCGCGTCCGCTCTCAGCTGTGAGCTGCGCTGGCGGACGTCATTAAGTATGCGTCC	1680
QY	1681	TGTGACAGATGTGTCTTCGCCGGGAGCCCATATTGCGGCTGGAGCGTCAACACCAAGCGCG	1740
Dp	1681	TGTGACAGATGTGTCTTCGCCGGGAGCCCATATTGCGGCTGGAGCGTCAACACCAAGCGCG	1740
QY	1741	TGTGTGGCGGTGGGTGGGCACTTTGGATCTTACTGATCAGCATGTGATACCTCGGAC	1800
Dp	1741	TGTGTGGCGGTGGGTGGGCACTTTGGATCTTACTGATCAGCATGTGATACCTCGGAC	1800
QY	1801	ACTTCAGGCACTTTCGAACCTCCGTGGCAGTAAATACAGTGAAGGCCCACTGCCAAAAC	1860
Dp	1801	ACTTCAGGCACTTTCGAACCTCCGTGGCAGTAAATACAGTGAAGGCCCACTGCCAAAAC	1860
QY	1861	ATCAACGGTGTGGCGGGGACAGACCTGGTGTGCTGCGTCCACTCTCTCCCAACTTGGCC	1920

Db	1861	ATCAGGTGTGGCGGCACAGACCTGGTGTCTGCCTGCACACCTCTCCAACTTGACC	19270
Qy	1921	CTGCCCGCACTCCAAACCCCGAGAGTTCATCATATGAGGGGAAACCCCAACCGCTCGGCGG	1980
Db	1921	CTGCCCGCACTCCAAACCCCGAGAGTTCATCATATGAGGGGAAACCCCAACCGCTCGGCGG	1980
Qy	1981	ANAGCGTGGAGGTGTAGCTCTACTTTTGACAGGCAACCAAGTATCTCAGGACATGGC	2040
Db	1981	ANAGCGTGGAGGTGTAGCTCTACTTTTGACAGGCAACCAAGTATCTCAGGACATGGC	2040
Qy	2041	ACGGGCACCTCTCTGTCTGGGACAGATATCTGGCCAGCACCACCCGGCCATGAGACCT	2100
Db	2041	ACGGGCACCTCTCTGTCTGGGACAGATATCTGGCCAGCACCACCCGGCCATGAGACCT	2100
Qy	2101	GCTGTGCTCAGACAGGGGACATGCACCTTGGTGGTCACTACAGGGGACCAAGCTCGGAGAAAG	2160
Db	2101	GCTGTGCTCAGACAGGGGACATGCACCTTGGTGGTCACTACAGGGGACCAAGCTCGGAGAAAG	2160
Qy	2161	CATCTCTCTCCTCTCTGTGTAATCAACAGACAGCGCGGACCCAGCGCCAAAATTTTCAAG	2220
Db	2161	CATCTCTCTCCTCTCTGTGTAATCAACAGACAGCGCGGACCCAGCGCCAAAATTTTCAAG	2220
Qy	2221	GCAGAGTTNAAAGATGTGTGTTGNTGATTTGACATGTGTGTTGTGTGTATGTG	2280
Db	2221	GCAGAGTTNAAAGATGTGTGTTGNTGATTTGACATGTGTGTTGTGTGTATGTG	2280
Qy	2281	TGTG 2284	
Db	2281	TGTG 2284	

```

RESULT 4
US-10-003-152-11
; Sequence 11, Application US/10003152
; Publication No. US20020151494A1
; GENERAL INFORMATION:
; APPLICANT: Shinketsu, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Yang, Meijia
; APPLICANT: Boldog, Ferenc
; APPLICANT: Herrmann, John
; TITLE OF INVENTION: No. US20020151494A1e1 Amino Acid Sequences for Human Semaphorin
; FILE REFERENCE: 15966-554 Cura-54 CON-S12
; CURRENT APPLICATION NUMBER: US/10/003.152
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (166)..(1935)
US-10-003-152-11

Query Match      82.6%; Score 1887.4; DB 13; Length 2155;
Best Local Similarity 99.5%; Pred. NO. 0;
Matches 1914; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

QY      1 CGGCCCTTCTACACTCCTCCCTGCTGATGTGGAACGGGGTTGGGGTTCTGCAGGGCT 60
        |||||||
DB      1 CGGCCCTTCTACACTCCTCCCTGCTGATGTGGAACGGGGTTGGGGTTCTGCAGGGCT 60

QY      61 ATTGTCTGCGTGGGGGAAGGGGACAGGCGCGGACCTCCGCTCGACGCCGCGC 120
        |||||||
DB      61 ATTGTCTGCGTGGGGGAAGGGGACAGGCGCGGACCTCCGCTCGACGCCGCGC 120
        |||||||

```

OY 121 ACCAGCAGACAGCTGGCCTTAAGCTCAGAGCCGGGCGCTGGCCATGAGCCCCACACCTGG 180
 DB 121 ACCAGCAGACAGCTGGCCTTAAGCTCAGAGCCGGGCGCTGGCCATGAGCCCCACACCTGG 180
 OY 181 GCTGTCTGGCTGGTGGGAGCAAGGCTGTGGGCGCTGGGCAATTGGGGCTGAGGTGGG 240
 DB 181 GCTGTCTGGCTGGTGGGAGCAAGGCTGTGGGCGCTGGGCAATTGGGGCTGAGGTGGG 240
 OY 241 AACCTTGTCCGCGTAAAGACAGTGTCTTCTGGGAGCTGGCCACAGTATGAGCGGCTTC 300
 DB 241 AACCTTGTCCGCGTAAAGACAGTGTCTTCTGGGAGCTGGCCACAGTATGAGCGGCTTC 300
 OY 301 TCCAGACCCGCGATCCAGAGACTTCTGACACTGACGCTGACGAGCCCACTGGGCTTTC 360
 DB 301 TCCAGACCCGCGATCCAGAGACTTCTGACACTGACGCTGACGAGCCCACTGGGCTTTC 360
 OY 361 TACGTTGGGCGCGGAGAGGCGCTGTTTCTTCAGCATGAGAGCCCTGGAGCTGCAAGGA 420
 DB 361 TACGTTGGGCGCGGAGAGGCGCTGTTTCTTCAGCATGAGAGCCCTGGAGCTGCAAGGA 420
 OY 421 GCGATCTCTGGAGAGGCGCGCTGGAGAAAGAGACTGATGTATCCAGAAAGGAGAAAC 480
 DB 421 GCGATCTCTGGAGAGGCGCGCTGGAGAAAGAGACTGATGTATCCAGAAAGGAGAAAC 480
 OY 481 AACCAAGCCGAGTGTCTTCACTTCACTCCGCTTCTGACAGCCCTACATGATCCCTCC 540
 DB 481 AACCAAGCCGAGTGTCTTCACTTCACTCCGCTTCTGACAGCCCTACATGATCCCTCC 540
 OY 541 TACGCTGTGGACACTGACGCTTCCAGGCGCAAGTGCATGACATGATGATGATGATG 600
 DB 541 TACGCTGTGGACACTGACGCTTCCAGGCGCAAGTGCATGACATGATGATGATGATG 600
 OY 601 ACTTGTGAGCATGAGAGATTGAAGATGGAAGGCAAGTGTCCCTATGACCCAGCTAAG 660
 DB 601 ACTTGTGAGCATGAGAGATTGAAGATGGAAGGCAAGTGTCCCTATGACCCAGCTAAG 660
 OY 661 GGCATCTGTGGCTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 720
 DB 661 GGCATCTGTGGCTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 720
 OY 721 GGCAGGAAACCATTAATCTGCTGATACATGAGGCGCCCACTCCATGATGATGATGATG 780
 DB 721 GGCAGGAAACCATTAATCTGCTGATACATGAGGCGCCCACTCCATGATGATGATGATG 780
 OY 781 CTGGGCTTGTGGCTTCAAGAACTCTACTTGTAGGCTTGGCTATGATGATGATGATG 840
 DB 781 CTGGGCTTGTGGCTTCAAGAACTCTACTTGTAGGCTTGGCTATGATGATGATGATG 840
 OY 841 GGCAGCTTCAAGGAGGAG 900
 DB 841 GGCAGCTTCAAGGAGGAG 900
 OY 901 TCCGACTGTATGCGGAG 960
 DB 901 TCCGACTGTATGCGGAG 960
 OY 961 GGGGCGGAG 1020
 DB 961 GGGGCGGAG 1020
 OY 1021 TCTGTCCCGGAACTGGACACTTACTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
 DB 1021 TCTGTCCCGGAACTGGACACTTACTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
 OY 1081 ACCTCTGGACAAACACACCTTCTTGGGGTTTTTCAAGACAGAGAGAGAGAGAGAGAG 1140
 DB 1081 ACCTCTGGACAAACACACCTTCTTGGGGTTTTTCAAGACAGAGAGAGAGAGAGAGAG 1140
 OY 1141 CTGTGCGGCAATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
 DB 1141 CTGTGCGGCAATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200

OY 1201 AAGAGTACCATGAG 1260
 DB 1201 AAGAGTACCATGAG 1260
 OY 1261 GGGCTGGCTGGTGGATTAACAACTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 DB 1261 GGGCTGGCTGGTGGATTAACAACTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 OY 1321 CTACCCGAGCAATCTCAACTTCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
 DB 1321 CTACCCGAGCAATCTCAACTTCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
 OY 1381 CCTGGTGGAG 1440
 DB 1381 CCTGGTGGAG 1440
 OY 1441 GACCGGTTTACAG 1500
 DB 1441 GACCGGTTTACAG 1500
 OY 1501 GGCCTGGCTTCAAG 1560
 DB 1501 GGCCTGGCTTCAAG 1560
 OY 1561 CTGTTTGAACAG 1620
 DB 1561 CTGTTTGAACAG 1620
 OY 1621 GCGGCTCCCGCTTCAAG 1680
 DB 1621 GCGGCTCCCGCTTCAAG 1680
 OY 1681 TGTGAG 1740
 DB 1681 TGTGAG 1740
 OY 1741 TGTGAG 1800
 DB 1741 TGTGAG 1800
 OY 1801 ACTTGAAGCATTTGCACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
 DB 1801 ACTTGAAGCATTTGCACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
 OY 1861 ATCAGGAG 1920
 DB 1861 ATCAGGAG 1920
 OY 1921 CTG 1923
 DB 1918 CGG 1920

RESULT 5
 US-10-002-050-11
 ; Sequence 11, Application US/10002050
 ; Publication No. US20030032095A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Shinketsu, Richard
 ; APPLICANT: Fernandes, Elma
 ; APPLICANT: Vernet, Corine
 ; APPLICANT: Yang, Meijia
 ; APPLICANT: Boldog, Ferenc
 ; APPLICANT: Hermann, John
 ; TITLE OF INVENTION: No. US20030032095A1 Nucleic Acid Sequences Encoding Human Se
 ; FILE REFERENCE: 15966-554 Cure-54 CON-514
 ; CURRENT FILING DATE: 2001-11-02
 ; PRIOR APPLICATION NUMBER: 09/604,286
 ; PRIOR FILING DATE: 2000-06-22
 ; PRIOR APPLICATION NUMBER: 60/140,584
 ; PRIOR FILING DATE: 1999-06-23
 ; NUMBER OF SEQ ID NOS: 49

SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (166)..(1935)
US-10-002-050-11

Query Match 82.6% Score 1887.4; DB 14; Length 2155;
Best Local Similarity 99.5% Pred. No. 0;
Matches 1914; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

OY 1 CGGCGCTTCTCACACTCCTGCTGATGTGGAAGCGGGTTGGGGTTCTGAGAGGCT 60
Db 1 CGGCGCTTCTCACACTCCTGCTGATGTGGAAGCGGGTTGGGGTTCTGAGAGGCT 60
OY 61 ATTGTGCGCTGGGGAAGGGGACAGCCGGGACCGGACCTCCGCTCGCAGCGGCGGC 120
Db 61 ATTGTGCGCTGGGGAAGGGGACAGCCGGGACCGGACCTCCGCTCGCAGCGGCGGC 120
OY 121 ACCAGCGAGACAGTGGCGCTGAAGCTCAGAGCCGGGGCGTGGCGCATGGCCCCACACTGG 180
Db 121 ACCAGCGAGACAGTGGCGCTGAAGCTCAGAGCCGGGGCGTGGCGCATGGCCCCACACTGG 180
OY 181 GCTGTCTGGCTGTGCGCAGCAAGGCTGTGGGGCTTGGGCGTGAAGGTGTGTGG 240
Db 181 GCTGTCTGGCTGTGCGCAGCAAGGCTGTGGGGCTTGGGCGTGAAGGTGTGTGG 240
OY 241 AACCTTGTGCGCGCTGAAGACAGTGTCTTCTGGGGAGCTGGCGACGAGTACGGCGGCTTC 300
Db 241 AACCTTGTGCGCGCTGAAGACAGTGTCTTCTGGGGAGCTGGCGACGAGTACGGCGGCTTC 300
OY 301 TCCGAGACCGGCACTCCAGAGCTTCGACGCTGACGCGGAGCGGACCCACCTGGGCTTCG 360
Db 301 TCCGAGACCGGCACTCCAGAGCTTCGACGCTGACGCGGAGCGGACCCACCTGGGCTTCG 360
OY 361 TACGTGGGCGCCGAGAGGCGCTTGTGCTTCAGATGAGGCGCTTGGAGCTCAAGGA 420
Db 361 TACGTGGGCGCCGAGAGGCGCTTGTGCTTCAGATGAGGCGCTTGGAGCTCAAGGA 420
OY 421 GCGATCTCTGCGGAGGCGCCCGTGGAGAGAACTGATGTATCCAGAAAGGAGAAAC 480
Db 421 GCGATCTCTGCGGAGGCGCCCGTGGAGAGAACTGATGTATCCAGAAAGGAGAAAC 480
OY 481 AACGAGACGAGTCTCAACTTCATCCGCTTCCTGAGAGCCCTCAATGCTCCCACTG 540
Db 481 AACGAGACGAGTCTCAACTTCATCCGCTTCCTGAGAGCCCTCAATGCTCCCACTG 540
OY 541 TACGTCTGTGGCAGCTACGCGCTTCAGGCCAAGTGAACCTTCAACATGCTCCTC 600
Db 541 TACGTCTGTGGCAGCTACGCGCTTCAGGCCAAGTGAACCTTCAACATGCTCCTC 600
OY 601 ACTTTGAGACATGAGAGATTGAAAGATGGAGGGAAGTGTCCCTATGACCCAGCTAAG 660
Db 601 ACTTTGAGACATGAGAGATTGAAAGATGGAGGGAAGTGTCCCTATGACCCAGCTAAG 660
OY 661 GGGCATGTGGCGCTTCTTGGAGTGTGAGCTGACTACGCGGCACAGCTCAACAACTTCTG 720
Db 661 GGGCATGTGGCGCTTCTTGGAGTGTGAGCTGACTACGCGGCACAGCTCAACAACTTCTG 720
OY 721 GGCAGGAAACCATTAATCTGCTACATGAGGGGCCCCACACTCATTAAGACAGAGTAC 780
Db 721 GGCAGGAAACCATTAATCTGCTACATGAGGGGCCCCACACTCATTAAGACAGAGTAC 780
OY 781 CTGGCGCTTTGGCTCAACGAACCTCACTTGTAGGCTGTGCTATGACTGAGAGTGTG 840
Db 781 CTGGCGCTTTGGCTCAACGAACCTCACTTGTAGGCTGTGCTATGACTGAGAGTGTG 840
OY 841 GGCAGCTTCAAGGGGGAGAGACAGAGTCACTTCTTCAAGGAGCGGGGAGTGTGAG 900
Db 841 GGCAGCTTCAAGGGGGAGAGACAGAGTCACTTCTTCAAGGAGCGGGGAGTGTGAG 900

OY 901 TCCGACTGCTATGCCAGACAGTGTGTGCTGTGCGCCCTGTCTGCAAGGGGATATG 960
Db 901 TCCGACTGCTATGCCAGACAGTGTGTGCTGTGCGCCCTGTCTGCAAGGGGATATG 960
OY 961 GGGGGCGACGAGCCTTCAGAGAGAGTGAACACAGTTCTTGAAGGGCGGCTGGCATGC 1020
Db 961 GGGGGCGACGAGCCTTCAGAGAGAGTGAACACAGTTCTTGAAGGGCGGCTGGCATGC 1020
OY 1021 TCTGCCCCGAACTGGCACTCTTACTTCAACAGAGTGGAGGATGACACCTTCAGAGAC 1080
Db 1021 TCTGCCCCGAACTGGCACTCTTACTTCAACAGAGTGGAGGATGACACCTTCAGAGAC 1080
OY 1081 ACCTCTGCGACACACACACCTCTTGGGGTTTTCAGAGCAGTGGGGGATGATGATGAC 1140
Db 1081 ACCTCTGCGACACACACACCTCTTGGGGTTTTCAGAGCAGTGGGGGATGATGATGAC 1140
OY 1141 CTGTGCGGCATCTGTGAGTACAGTTGGAGAGATCCAGCGGCTGTGAGGGCCCTAT 1200
Db 1141 CTGTGCGGCATCTGTGAGTACAGTTGGAGAGATCCAGCGGCTGTGAGGGCCCTAT 1200
OY 1201 AAGGAGTACCATGAGAGAGCCGAGAGGAGCCGCTACACTGACCCCTGACAGCCCT 1260
Db 1201 AAGGAGTACCATGAGAGAGCCGAGAGGAGCCGCTACACTGACCCCTGACAGCCCT 1260
OY 1261 CGGCGTGGCTGTGCTATTACAGAGTGGGCGCCAGCGCTACACAGCTCCCTGGAG 1320
Db 1261 CGGCGTGGCTGTGCTATTACAGAGTGGGCGCCAGCGCTACACAGCTCCCTGGAG 1320
OY 1321 CTACCCGACAAATCTCTCACTGATCAGAGAGCCAGCCGCTGATGAGAGAGAGTGGGG 1380
Db 1321 CTACCCGACAAATCTCTCACTGATCAGAGAGCCAGCCGCTGATGAGAGAGAGTGGGG 1380
OY 1381 CCTGCGTGGAGAGCCCGCTGCTGTAAGAGAGGAGGAGCACTTACCACCTGGTGGCC 1440
Db 1381 CCTGCGTGGAGAGCCCGCTGCTGTAAGAGAGGAGGAGCACTTACCACCTGGTGGCC 1440
OY 1441 GACCGGGTTACAGAGACTTGTATGAGAGCCAGCTATAGTGTCTTCAATGGCAGAGAGAC 1500
Db 1441 GACCGGGTTACAGAGACTTGTATGAGAGCCAGCTATAGTGTCTTCAATGGCAGAGAGAC 1500
OY 1501 GCGTGGCTGTCAAGGCTGTGAGGCTGGGGCGCTGGGTTCACTGATGAGAGAGCTGCAG 1560
Db 1501 GCGTGGCTGTCAAGGCTGTGAGGCTGGGGCGCTGGGTTCACTGATGAGAGAGCTGCAG 1560
OY 1561 CTGTTTGACAGAGAGCCCATGAGAGGCTGTGATCTGAGAGCAAGACCTGCTCTT 1620
Db 1561 CTGTTTGACAGAGAGCCCATGAGAGGCTGTGATCTGAGAGCAAGACCTGCTCTT 1620
OY 1621 GCGGCTCCCGCTCTCAGAGCTGTGAGGCTGGCGGCGAGCTGATTAATATGCTGCTC 1680
Db 1621 GCGGCTCCCGCTCTCAGAGCTGTGAGGCTGGCGGCGAGCTGATTAATATGCTGCTC 1680
OY 1681 TGTGACAGACTGTCTGCTGCGCCGAGCCCTATTTGCGCTGAGCGCTCAACAGAGCCGC 1740
Db 1681 TGTGACAGACTGTCTGCTGCGCCGAGCCCTATTTGCGCTGAGCGCTCAACAGAGCCGC 1740
OY 1741 TGTGAGCGCGTGGGGGAGCTTGGATCTTACATGACAGAGTGTGATGAGTACCTGGAG 1800
Db 1741 TGTGAGCGCGTGGGGGAGCTTGGATCTTACATGACAGAGTGTGATGAGTACCTGGAG 1800
OY 1801 ACTTCAGGCAATTTGCAACCTCCGTGGCAGTAAAGTACAGTACAGGCGCCACTACCAAAAC 1860
Db 1801 ACTTCAGGCAATTTGCAACCTCCGTGGCAGTAAAGTACAGTACAGGCGCCACTACCAAAAC 1860
OY 1861 ATCAGGCTGTGGGGGAGAGAGAGCTGTGCTGCGCTGCGACCTCTCTCAACTTGGCC 1920
Db 1861 ATCAGGCTGTGGGGGAGAGAGAGCTGTGCTGCGCTGCGACCTCTCTCAACTTGGCC 1920
OY 1921 CTG 1923 1917
Db 1918 CGG-1920


```
RESULT 6
US-10-002-304-11
/ Sequence 11, Application US/10002304
/ Publication No. US20030036185A1
/ GENERAL INFORMATION:
/ APPLICANT: Shinkabe, Richard
/ APPLICANT: Fernandes, Elma
/ APPLICANT: Vernet, Corine
/ APPLICANT: Yang, Meljia
/ APPLICANT: Boldoy, Forenc
/ APPLICANT: Herrmann, John
/ TITLE OF INVENTION: Polynucleotides and polypeptides encoded thereby
/ FILE REFERENCE: 13366-354 Cura-34 CON-58
/ CURRENT APPLICATION NUMBER: US/10/002,304
/ PRIOR FILING DATE: 2001-11-02
/ PRIOR APPLICATION NUMBER: 09/604,286
/ PRIOR FILING DATE: 2000-06-22
/ PRIOR APPLICATION NUMBER: 60/140,584
/ NUMBER OF SEQ ID NOS: 49
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 11
/ LENGTH: 2155
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (166)..(1935)
US-10-002-304-11

Query Match      82.6% Score 1887.4; DB 14; Length 2155;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1914; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

QY      1 CGGCCCTCTCACTCCGCGCTGATGTGGAACGGGGTTGGGGTTGCGAGGCT 60
DB      1 CGGCCCTCTCACTCCGCGCTGATGTGGAACGGGGTTGGGGTTGCGAGGCT 60
QY      61 ATTGTCTGCGTGGGGAAGGAGGAGCGCGGAGCCGCGGACTCCGCTGCGACCGCGCG 120
DB      61 ATTGTCTGCGTGGGGAAGGAGGAGCGCGGAGCCGCGGACTCCGCTGCGACCGCGCG 120
QY      121 ACCAGCAGGACAGCTGGGCTGAAAGTCAAGCGCGGGCGTGGCCATGCGCCACACTGG 180
DB      121 ACCAGCAGGACAGCTGGGCTGAAAGTCAAGCGCGGGCGTGGCCATGCGCCACACTGG 180
QY      181 GCTGTCTGGCTGCTGGCAGCAGAGGCTGTGGGGCTGGGCAATTGGGGCTGAGGTGTGG 240
DB      181 GCTGTCTGGCTGCTGGCAGCAGAGGCTGTGGGGCTGGGCAATTGGGGCTGAGGTGTGG 240
QY      241 AACCTTGTGCGCGCTAAGACAGTGTCTTGGGGAGCTGGCCAGCGTACTACGGCGCTTC 300
DB      241 AACCTTGTGCGCGCTAAGACAGTGTCTTGGGGAGCTGGCCAGCGTACTACGGCGCTTC 300
QY      301 TCCGAGACCGGAGATCCAGAGACTTCTGACACTGACGCTGACGAGGCCACTGGCTTCTG 360
DB      301 TCCGAGACCGGAGATCCAGAGACTTCTGACACTGACGCTGACGAGGCCACTGGCTTCTG 360
QY      361 TACGATGGGGCGCCGAGAGGCGCTTGTGCTTACAGATGAGAGCGCTGAGAGTGAAGA 420
DB      361 TACGATGGGGCGCCGAGAGGCGCTTGTGCTTACAGATGAGAGCGCTGAGAGTGAAGA 420
QY      421 GCGATCTCTCTGGAGAGCCCGCTGGAGAGAAAGACTGATGTATCCAGAAAGGAAAGAC 480
DB      421 GCGATCTCTCTGGAGAGCCCGCTGGAGAGAAAGACTGATGTATCCAGAAAGGAAAGAC 480
QY      481 AACCAAGACCGAGTGTCTCAACTTCAATCCGCTTCTGACAGCCCTACAAATGCTCCAC 540
DB      481 AACCAAGACCGAGTGTCTCAACTTCAATCCGCTTCTGACAGCCCTACAAATGCTCCAC 540
QY      541 TACGCTGTGGGAGCTACGCGCTTCCAGAGCCAAAGTGCACCTACGTCAACATGCTCAC 600
DB      541 TACGCTGTGGGAGCTACGCGCTTCCAGAGCCAAAGTGCACCTACGTCAACATGCTCAC 600
```

```
DB      541 TACGCTGTGGGAGCTACGCGCTTCCAGAGCCAAAGTGCACCTACGTCAACATGCTCAC 600
QY      601 ACTTTGAGCATGAGAGATTGAAGATGGAAGGCAAGTGTCCCTATGACCACGCTAAG 660
DB      601 ACTTTGAGCATGAGAGATTGAAGATGGAAGGCAAGTGTCCCTATGACCACGCTAAG 660
QY      661 GGCATGCTGCGCTTCTTGTGGATGTGAGCTGTACTCGGCCACACTCAACAACTTCTCG 720
DB      661 GGCATGCTGCGCTTCTTGTGGATGTGAGCTGTACTCGGCCACACTCAACAACTTCTCG 720
QY      721 GGCAGGAAACCATATATCTGCTGAACATGGGGGCCCCACCACTCCATGAAAGCAAGTAC 780
DB      721 GGCAGGAAACCATATATCTGCTGAACATGGGGGCCCCACCACTCCATGAAAGCAAGTAC 780
QY      781 CTGGCTTTTGGCTCAACGAACTCCTGAGCTGCTGCTATGACTAGAGTGTG 840
DB      781 CTGGCTTTTGGCTCAACGAACTCCTGAGCTGCTGCTATGACTAGAGTGTG 840
QY      841 GGCACCTTCACGGGGAGCAGACAGCAAGCTCTACTTCTTCTTCAGGAGCGGCACTGGAG 900
DB      841 GGCACCTTCACGGGGAGCAGACAGCAAGCTCTACTTCTTCTTCAGGAGCGGCACTGGAG 900
QY      901 TCCGACTGCTATGCGGAGAGGTGGTGGCTGCTGGCCCGCCGCTGCAAGGGCGATATG 960
DB      901 TCCGACTGCTATGCGGAGAGGTGGTGGCTGCTGGCCCGCCGCTGCAAGGGCGATATG 960
QY      961 GCGGCGCAGCAGACCTCGAGAGAGAAAGTGAACAGCTTCTGTAAGCGCGCGCTGCAATG 1020
DB      961 GCGGCGCAGCAGACCTCGAGAGAGAAAGTGAACAGCTTCTGTAAGCGCGCGCTGCAATG 1020
QY      1021 TCTGCCCGCACTGCGAGCTCTACTTCAACAGCTGAGCGCATCACACCTCGAGAC 1080
DB      1021 TCTGCCCGCACTGCGAGCTCTACTTCAACAGCTGAGCGCATCACACCTCGAGAC 1080
QY      1081 ACCTCTGGCAACAACACGCTCTTTGGGGTTTCAAGCAACAGTGGGCTGACATGTAC 1140
DB      1081 ACCTCTGGCAACAACACGCTCTTTGGGGTTTCAAGCAACAGTGGGCTGACATGTAC 1140
QY      1141 CTGTGCGCATCTGTGAGTACAGCTTGGAAAGATCAACAGCGGGTGTGGAGGCCCTAT 1200
DB      1141 CTGTGCGCATCTGTGAGTACAGCTTGGAAAGATCAACAGCGGGTGTGGAGGCCCTAT 1200
QY      1201 AAGGATACATGAGAGAAAGCCAGAAAGTGGACCGCTACACTGACCTGTACCCAGCCT 1260
DB      1201 AAGGATACATGAGAGAAAGCCAGAAAGTGGACCGCTACACTGACCTGTACCCAGCCT 1260
QY      1261 GCGCTGGCTGTGCAATTAACACTGGCATGGCGGCAAGGGCTACACAGCTCCCTGGAG 1320
DB      1261 GCGCTGGCTGTGCAATTAACACTGGCATGGCGGCAAGGGCTACACAGCTCCCTGGAG 1320
QY      1321 CTACCGACACACATCTCAACTTGTCAAGAAAGACACCGCTGATGAGAGCAGGTGGAG 1380
DB      1321 CTACCGACACACATCTCAACTTGTCAAGAAAGACACCGCTGATGAGAGCAGGTGGAG 1380
QY      1381 CTTGCGTGGAGCGCGCCCTGCTGTGTAAGAAAGGCAACCACTTCAACCACTGGTGGC 1440
DB      1381 CTTGCGTGGAGCGCGCCCTGCTGTGTAAGAAAGGCAACCACTTCAACCACTGGTGGC 1440
QY      1441 GACCGGTTACAGAGCTGATGAGAGCACCTTATAGTCTTCAATGGGCAAGAGAC 1500
DB      1441 GACCGGTTACAGAGCTGATGAGAGCACCTTATAGTCTTCAATGGGCAAGAGAC 1500
QY      1501 GCGTGGCTGCTCAAGGCTGTGAGCTGGGGCCCTGGTTCACCTGATTTGAGAGAGTGCAG 1560
DB      1501 GCGTGGCTGCTCAAGGCTGTGAGCTGGGGCCCTGGTTCACCTGATTTGAGAGAGTGCAG 1560
QY      1561 CTGTTTGAACAGAGCCCATGAGAAAGCTGTGCTATCTCAGAGCAAGAGTGTCTTT 1620
DB      1561 CTGTTTGAACAGAGCCCATGAGAAAGCTGTGCTATCTCAGAGCAAGAGTGTCTTT 1620
QY      1621 GCGGCGCTCCGCTCTCAGTGTGAGAGTGGCGCGCGGCACTGATTAAGTATGCTTC 1680
DB      1621 GCGGCGCTCCGCTCTCAGTGTGAGAGTGGCGCGCGGCACTGATTAAGTATGCTTC 1680
```



```

OY 1321 CTACCCGACACATCTCTCACTTCTGTCAGAGAACCCGCTGATGAGAGACAGGTGGG 1380
    |||||||
DB 1321 CTACCCGACACATCTCTCACTTCTGTCAGAGAACCCGCTGATGAGAGACAGGTGGG 1380
OY 1381 CCTGGGTGAGACCCCGCTCTCTGTGAGAGAGGACCAACTTACCCACCTGCTGGCC 1440
    |||||||
DB 1381 CCTGGGTGAGACCCCGCTCTCTGTGAGAGAGGACCAACTTACCCACCTGCTGGCC 1440
OY 1441 GACCGGTTACAGAGACTGATGAGAGCCACTATACAGTGTCTTCAATTGGACAGAGAC 1500
    |||||||
DB 1441 GACCGGTTACAGAGACTGATGAGAGCCACTATACAGTGTCTTCAATTGGACAGAGAC 1500
OY 1501 GCGTGGCTGCTCAAGAGCTGTGAGCCCTGGGCCCTGGGTTCACTGATGAGAGCTGGAG 1560
    |||||||
DB 1501 GCGTGGCTGCTCAAGAGCTGTGAGCCCTGGGCCCTGGGTTCACTGATGAGAGCTGGAG 1560
OY 1561 CTGTTTACAGAGAGCCCATAGAGAGCTGTCTATCTCAGAGAGCAAGACCTGCTTT 1620
    |||||||
DB 1561 CTGTTTACAGAGAGCCCATAGAGAGCTGTCTATCTCAGAGAGCAAGACCTGCTTT 1620
OY 1621 GCCGGCTCCGCTCTCAAGCTGTGAGCTGGCCGCTGGCCGACTGATTAAGTATCGCTCC 1680
    |||||||
DB 1621 GCCGGCTCCGCTCTCAAGCTGTGAGCTGGCCGCTGGCCGACTGATTAAGTATCGCTCC 1680
OY 1681 TGTGACAGACTGTGTCTCTGCGCGGAGCCCTATTGGCGCTGAGCGTCAACAGCCGCG 1740
    |||||||
DB 1681 TGTGACAGACTGTGTCTCTGCGCGGAGCCCTATTGGCGCTGAGCGTCAACAGCCGCG 1740
OY 1741 TGTGTGGCGCTGGGTGGGCACTTTGAGACTTACTGATCCAGAGATGATGATGATGATGAT 1800
    |||||||
DB 1741 TGTGTGGCGCTGGGTGGGCACTTTGAGACTTACTGATCCAGAGATGATGATGATGATGAT 1800
OY 1801 ACTTCAGGACATTTTCAACCTCTCTGTCAGTAAAGATACAGAGCCCACTTCCAAAAAC 1860
    |||||||
DB 1801 ACTTCAGGACATTTTCAACCTCTCTGTCAGTAAAGATACAGAGCCCACTTCCAAAAAC 1860
OY 1861 ATTCAGGCTGTGGGGGACAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
    |||||||
DB 1861 ATTCAGGCTGTGGGGGACAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
OY 1921 CTG 1923
    ||
DB 1918 CGG 1920

```

RESULT 8
US-10-002-050-21
Sequence 21, Application US/10002050
Publication No. US20030032095A1

GENERAL INFORMATION:
APPLICANT: Shinkels, Richard

APPLICANT: Fernandez, Elma

APPLICANT: Vernet, Corine

APPLICANT: Yang, Melja

APPLICANT: Boldog, Ferenc

APPLICANT: Heitmann, John

TITLE OF INVENTION: No. US20030032095A1el Nucleic Acid Sequences Encoding Human Sema

FILE REFERENCE: 15966-534 Cura-54 CON-514

CURRENT FILING DATE: US/10/002, 050

PRIOR FILING DATE: 2001-11-02

PRIOR APPLICATION NUMBER: 09/604, 286

PRIOR FILING DATE: 2000-06-22

PRIOR APPLICATION NUMBER: 60/140, 584

PRIOR FILING DATE: 1999-06-23

NUMBER OF SEQ ID NOS: 49

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 21

LENGTH: 2156

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (166)..(2037)
NAME/KEY: variation
LOCATION: (1)..(2156)
OTHER INFORMATION: N may be any nucleotide
US-10-002-050-21

Query Match 82.3%; Score 1879.4; DB 14; Length 2156;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1909; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

```

OY 1 CGGCCCTTCACACTCTCGCCCTGCTATGTAAGAGGGGTTGGGGTTCTGACAGGCT 60
    |||||||
DB 1 CGGCCCTTCACACTCTCGCCCTGCTATGTAAGAGGGGTTGGGGTTCTGACAGGCT 60
OY 61 ATTGTCTGCTGGGGAAGGGAGAGCCGAGACCTGCGCTGCGACGCGCGCC 120
    |||||||
DB 61 ATTGTCTGCTGGGGAAGGGAGAGCCGAGACCTGCGCTGCGACGCGCGCC 120
OY 121 ACCAGCAGACAGCTGGCCTGAAGCTCAGAGCCGGGGCGTGCAGATGGCCCACTGG 180
    |||||||
DB 121 ACCAGCAGACAGCTGGCCTGAAGCTCAGAGCCGGGGCGTGCAGATGGCCCACTGG 180
OY 181 GCTGTCTGGCTGTGGAGCAAGGCTGTGGGGCGCTGGGCAATTGGGGCTGAGTGTGG 240
    |||||||
DB 181 GCTGTCTGGCTGTGGAGCAAGGCTGTGGGGCGCTGGGCAATTGGGGCTGAGTGTGG 240
OY 241 AACCTTGTGCGCGTAAGACAGTGTCTTCTGGGAGCTGGCCAGAGTATGAGCGGCTTC 300
    |||||||
DB 241 AACCTTGTGCGCGTAAGACAGTGTCTTCTGGGAGCTGGCCAGAGTATGAGCGGCTTC 300
OY 301 TCCAGACCGGCAATCCAGAGACTTCTGACACTGACGCTGACGAGAGCCACTGGGCTTCG 360
    |||||||
DB 301 TCCAGACCGGCAATCCAGAGACTTCTGACACTGACGCTGACGAGAGCCACTGGGCTTCG 360
OY 361 TACGTTGGGCGCCGAGAGAGCCCTGTTGCTTCAAGATGAGAGGCGCTGAGCTGCAAGA 420
    |||||||
DB 361 TACGTTGGGCGCCGAGAGAGCCCTGTTGCTTCAAGATGAGAGGCGCTGAGCTGCAAGA 420
OY 421 GCGATCTCTGAGAGAGCCCGCTGAGAGAGAGAGTGTATCCAGAGAGGAGAAC 480
    |||||||
DB 421 GCGATCTCTGAGAGAGCCCGCTGAGAGAGAGAGTGTATCCAGAGAGGAGAAC 480
OY 481 AACGAGACGAGTGTCTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
    |||||||
DB 481 AACGAGACGAGTGTCTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
OY 541 TACGCTGTGACACTTACGCTTCAGAGCCAGAGAGTGCATGATGATGATGATGATGATGAT 600
    |||||||
DB 541 TACGCTGTGACACTTACGCTTCAGAGCCAGAGAGTGCATGATGATGATGATGATGATGAT 600
OY 601 ACTTTGAGCATGAGAGATTGAAGATGGAAGGCAAGTGTCCCTATGACCCAGCTAAG 660
    |||||||
DB 601 ACTTTGAGCATGAGAGATTGAAGATGGAAGGCAAGTGTCCCTATGACCCAGCTAAG 660
OY 661 GCGCATGTGCTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
    |||||||
DB 661 GCGCATGTGCTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
OY 721 GGCAGGAAACCATTTATCTGCTGTAACATGAGGCGCCACCACTCATGAAACAGAGTAC 780
    |||||||
DB 721 GGCAGGAAACCATTTATCTGCTGTAACATGAGGCGCCACCACTCATGAAACAGAGTAC 780
OY 781 CTGGCTTTTGGCTCAAGCACTTCTGAGGCTGTGCTTATGCTGAGAGAGTGTG 840
    |||||||
DB 781 CTGGCTTTTGGCTCAAGCACTTCTGAGGCTGTGCTTATGCTGAGAGAGTGTG 840
OY 841 GGCAGCTTCAGAGGAGAGAGAGAGAGTCTTCTTCTTCAAGGAGAGGAGAGTGTGAG 900
    |||||||
DB 841 GGCAGCTTCAGAGGAGAGAGAGAGAGTCTTCTTCTTCAAGGAGAGGAGAGTGTGAG 900
OY 901 TCCGACTGATGCGAGAGAGTGTGTGCTGTGCGCCGCTGTCTGCAAGGCGCATATG 960
    |||||||
DB 901 TCCGACTGATGCGAGAGAGTGTGTGCTGTGCGCCGCTGTCTGCAAGGCGCATATG 960

```


QY 601 ACTTGAGCATGAGAGATTGAAAGATGGAAAGGCAAGTGTCCCTATGACCCAGCTAAG 660
 DB 601 ACTTGGAGCATGAGAGATTGAAAGATGGAAAGGCAAGTGTCCCTATGACCCAGCTAAG 660
 QY 661 GGGCATTGTCGTCCTTCTGTGATGATGAGTGTACTGTGGCCACACTCAACAACCTTCTG 720
 DB 661 GGGCATTGTCGTCCTTCTGTGATGATGAGTGTACTGTGGCCACACTCAACAACCTTCTG 720
 QY 721 GGCACGGAACCAATTATCTGTGCTAACTATGGGGCCCACTCCATGATGAAGACAGTAC 780
 DB 721 GGCACGGAACCAATTATCTGTGCTAACTATGGGGCCCACTCCATGATGAAGACAGTAC 780
 QY 781 CTGGCTTTTGGCTTAACGAACCTTCTGAGCTTGTGCTATGTACCTGAGAGTGTG 840
 DB 781 CTGGCTTTTGGCTTAACGAACCTTCTGAGCTTGTGCTATGTACCTGAGAGTGTG 840
 QY 841 GGCACCTTCAAGGGGAGAGAGACAAAGGCTACTCTTCTTCAGGAGAGGGGCACTGGAG 900
 DB 841 GGCACCTTCAAGGGGAGAGAGACAAAGGCTACTCTTCTTCAGGAGAGGGGCACTGGAG 900
 QY 901 TCCGACTGCTATGCGAGCAGAGTGGTGGCTGTGGCCGCTGTCTGCAAGGGCGATATG 960
 DB 901 TCCGACTGCTATGCGAGCAGAGTGGTGGCTGTGGCCGCTGTCTGCAAGGGCGATATG 960
 QY 961 GGGGGCGACGAGACCTTGCAGAGAGTGGACACAGTTCCTGAAGGGCGGCTGGCATGC 1020
 DB 961 GGGGGCGACGAGACCTTGCAGAGAGTGGACACAGTTCCTGAAGGGCGGCTGGCATGC 1020
 QY 1021 TCTGCCCGCAACTGAGCTGCTACTTCAACGAGCTGAGGGGATGACACACCTGAGAGAC 1080
 DB 1021 TCTGCCCGCAACTGAGCTGCTACTTCAACGAGCTGAGGGGATGACACACCTGAGAGAC 1080
 QY 1081 ACCTCTGGCAACACACACCTTCTTGGGGTTTTTCAAGCAGACTGGGGTGCATGTAC 1140
 DB 1081 ACCTCTGGCAACACACACCTTCTTGGGGTTTTTCAAGCAGACTGGGGTGCATGTAC 1140
 QY 1141 CTGTGGGCTATCTGTGATGACAGAGTGGAAAGATTCACAGGGGTGTTGAGGGCCCTAT 1200
 DB 1141 CTGTGGGCTATCTGTGATGACAGAGTGGAAAGATTCACAGGGGTGTTGAGGGCCCTAT 1200
 QY 1201 AAGGAGTACATGAG 1260
 DB 1201 AAGGAGTACATGAG 1260
 QY 1261 GGGCCTGCTGCTGATTAACAACTGGCATGGCGCAGCGGCTACACAGCTCCCTGGAG 1320
 DB 1261 GGGCCTGCTGCTGATTAACAACTGGCATGGCGCAGCGGCTACACAGCTCCCTGGAG 1320
 QY 1321 CTACCCGCAACATCTCTCACTTGTCAAGAAAGCAGCCGCTGATGAGAGAGAGTGGG 1380
 DB 1321 CTACCCGCAACATCTCTCACTTGTCAAGAAAGCAGCCGCTGATGAGAGAGAGTGGG 1380
 QY 1381 CCTGCGTGGAG 1440
 DB 1381 CCTGCGTGGAG 1440
 QY 1441 GACCGGTTTACAGAGACTTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
 DB 1441 GACCGGTTTACAGAGACTTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
 QY 1501 GGGTGGCTGCTCAAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
 DB 1501 GGGTGGCTGCTCAAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
 QY 1561 CTGTTTGAACAG 1620
 DB 1561 CTGTTTGAACAG 1620
 QY 1621 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
 DB 1621 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680

QY 1681 TGTGACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
 DB 1681 TGTGACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
 QY 1741 TGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
 DB 1741 TGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
 QY 1801 ACTTACGAGCTTGTGACACTTCCGTCAGTAAAGATACAGTACAGGCGCACTNCCCAAAAC 1860
 DB 1801 ACTTACGAGCTTGTGACACTTCCGTCAGTAAAGATACAGTACAGGCGCACTNCCCAAAAC 1860
 QY 1861 ATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
 DB 1861 ATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
 QY 1921 CTG 1923
 DB 1918 CCG 1920

RESULT 10
 US-10-149-819-25
 ; Sequence 25, Application US/10149819
 ; Publication No. US20030044913A1
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE GENOMICS, INC.
 ; APPLICANT: YUE, Henry
 ; APPLICANT: AZIMZAI, Valda
 ; APPLICANT: TANG, Y. Tom
 ; APPLICANT: PATTERSON, Chandra
 ; APPLICANT: BAUGHN, Mariah R.
 ; APPLICANT: SHAH, Purva
 ; APPLICANT: LAU, Preeti
 ; APPLICANT: AU-Young, Janice
 ; APPLICANT: BURFORD, Neil
 ; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
 ; FILE REFERENCE: PF-0760 PCT
 ; CURRENT FILING DATE: US/10/149,819
 ; PRIOR APPLICATION NUMBER: 2002-06-10
 ; PRIOR FILING DATE: 60/172,852; 60/172,354
 ; NUMBER OF SEQ ID NOS: 42
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 25
 ; LENGTH: 3293
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No. US20030044913A1 1674661CB1
 US-10-149-819-25

Query Match 77.1%; Score 1761; DB 14; Length 3293;
 Best Local Similarity 97.4%; Pred. No. 0;
 Matches 1811; Conservative 0; Mismatches 45; Indels 3; Gaps 2;
 QY 84 CAGGCGGAGACGGGACGCTCCGTCGAGCGCGCGGACACACAGAGAGAGAGAGAGAGAG 143
 DB 52 CCGAAGCGGAGAGGCGCGCGCGCGGAGCTGAGACTGCGCGCAATTAAGAGAGAGAGAGAG 111
 QY 144 GCTCAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 203
 DB 112 GCTCAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 171
 QY 204 GCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 263
 DB 172 GCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 231
 QY 264 GTCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 323
 DB 232 GTCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 291

QY 382 CCGTCAACATGCTACCTTCTGAGCAGAGAGTTGAAGTGGGAGCGCAAGTG 641
DB 6007 CCGTCAACATGCTACCTTCTGAGCAGAGAGTTGAAGTGGGAGCGCAAGTG 6066
QY 642 TCCCATGAGCCAGCTAAGGGGCGATGCGCTTCTTGGATGGTGGAGCTACTCGGC 701
DB 6067 TCCCATGAGCCAGCTAAGGGGCGATGCGCTTCTTGGATGGTGGAGCTACTCGGC 6126
QY 702 CACACTCAACAACTTCTGCGGACCGAACCCATATCCGCTAACATGAGGGGCCACCA 761
DB 6127 CACACTCAACAACTTCTGCGGACCGAACCCATATCCGCTAACATGAGGGGCCACCA 6186
QY 762 CTCCATGAAGAGAGATGATGCTGGCTTTTGGCTCAACGAGACCTGATGAGCTGAC 821
DB 6187 CTCCATGAAGAGAGATGATGCTGGCTTTTGGCTCAACGAGACCTGATGAGCTGAC 6246
QY 822 CTATGATGATGAGATGATGAGGAGCTTACGAGGGGAGCGAGCAAGGCTTACTTCTG 881
DB 6247 CTATGATGATGAGATGATGAGGAGCTTACGAGGGGAGCGAGCAAGGCTTACTTCTG 6306
QY 882 CAGGAGCGGGGAGTGGAGTCCGACTGCTATGCTCCGAGCAGAGTGTGGCTGCTGAGCGG 941
DB 6307 CAGGAGCGGGGAGTGGAGTCCGACTGCTATGCTCCGAGCAGAGTGTGGCTGCTGAGCGG 6366
QY 942 TGTCTGCAAGGGGCGATATGGGGGGCGACAGCCTGCGAGAGAGAGTGGACAGCTTCT 1001
DB 6367 TGTCTGCAAGGGGCGATATGGGGGGCGACAGCCTGCGAGAGAGAGTGGACAGCTTCT 6426
QY 1002 GAAGCGGGGCGGATGCTGCTGCGGCGAAGTGGAGCTTCAACAGCAGCTGAGCG 1061
DB 6427 GAAGCGGGGCGGATGCTGCTGCGGCGAAGTGGAGCTTCAACAGCAGCTGAGCG 6486
QY 1062 GATGCAACACCTGAGAGACCTCTGCGACAAACACACCTTCTTGGGGTTTTCAAGC 1121
DB 6487 GATGCAACACCTGAGAGACCTCTGCGACAAACACACCTTCTTGGGGTTTTCAAGC 6546
QY 1122 ACAGGGGGGTGACATGATGCTGCGGCGATGCTGAGTACAGTGGAGAGATGAGCG 1181
DB 6547 ACAGGGGGGTGACATGATGCTGCGGCGATGCTGAGTACAGTGGAGAGATGAGCG 6606
QY 1182 GGTGTTTGAAGGGGCGCTTAAAGAGATGAGTACAGTGGAGAGAGTGGAGAGTGGAG 1241
DB 6607 GGTGTTTGAAGGGGCGCTTAAAGAGATGAGTACAGTGGAGAGAGTGGAGAGTGGAG 6666
QY 1242 TGACCTGTACCCAGCCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1301
DB 6667 TGACCTGTACCCAGCCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6726
QY 1302 CTACACAGCTCCCTGAGGCTACCCGAGACATCTCTCAAACTTCTGCAAGAGACCCGCT 1361
DB 6727 CTACACAGCTCCCTGAGGCTACCCGAGACATCTCTCAAACTTCTGCAAGAGACCCGCT 6786
QY 1362 GATGAGAGAGAGTGGGGGCGCTGCTGAGAGCGGCGGCTGCTGCTGCTGCTGCTGCTG 1421
DB 6787 GATGAGAGAGAGTGGGGGCGCTGCTGAGAGCGGCGGCTGCTGCTGCTGCTGCTGCTG 6846
QY 1422 CTTCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1481
DB 6847 CTTCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6906
QY 1482 GTTCATGAGAGAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1541
DB 6907 GTTCATGAGAGAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6966
QY 1542 CCGTATGAGAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1601
DB 6967 CCGTATGAGAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7026
QY 1602 GAGCAAG--AAGCTGCTCTTGGCGGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1658
DB 7027 GAGCAAGTAAAGCTGCTCTTGGCGGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7086
QY 1659 GAGCTGATAAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1718

DB 7087 CCACTGCAATGAAGTATCGCTCTGTCAGACATGTCCTGCGCGGAGCCCTATTCGCG 7146
QY 1719 CCGAGCGGCAACACAGCGGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1778
DB 7147 CCGAGCGGCAACACAGCGGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7206
QY 1779 CCGAGTGTGATGACCTCGGACATTCAGGCAATTCGCAACCTCCGTCGAGTAAAGTACA 1838
DB 7207 CCGAGTGTGATGACCTCGGACATTCAGGCAATTCGCAACCTCCGTCGAGTAAAG--AAA 7264
QY 1839 CTCAGCGGCACTNCCAAAACATCACGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGG 1898
DB 7265 CTCAGCGGCACT--CCAAAACATCACGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGG 7323
QY 1899 CCACCTCTCTCCAACTTGGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1942
DB 7324 CCACCTCTCTCCAACTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7367

RESULT 12
US-09-759-130B-336
; Sequence 336, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kistl, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodenart, Andrew
; APPLICANT: Holzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; FILE REFERENCE: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE REFERENCE: MP100-5350WIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1998-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 336
; LENGTH: 3503
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-759-130B-336
Query Match 59.2%; Score 1351; DB 11; Length 3503;
Best Local Similarity 84.7%; Pred. No. 0;


```

Db      1505  CGCCGCGCTGCGGGAAGAGCTGAGAAAGGGGCGCAAGGCTACAGAGGAGCT--TGG 1562
Oy      1550  AGGAGCTGAGCTGTTTGACACAGAGCCATGAGAAGCGTGGTGTATCTCAGACAAAGA 1609
Db      1563  TGTACCCCTGAGCTGCGCAAGAGAGCCACAGTCCCTCC--GGCCCTGCTCTGA 1620
Oy      1610  AGCTGCTCTTGGCCGCTCCCGCTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1669
Db      1621  ACCAATATGAAACCTTGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
Oy      1670  AGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1729
Db      1681  AGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Oy      1730  ACACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1789
Db      1741  CCCAGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Oy      1790  TGACCTGCGACCTTCAAGCAATTTGCAACCTCCGCTGCGAGTAAATACAGTACAGCCAC 1849
Db      1801  TGCCAAATGCTTACGCTGCTTACAACTAGGAGGAGAGCCGCGGAGGCTGCGGACCC 1860
Oy      1850  TNCCTCAAAACATCAAGGTGCTGCGGCGACAGACCTGCTGCTGCTGCTGCTGCTGCTGCT 1909
Db      1861  CCTGCTGAGCTGCTGCGGATGAACCTAGAGAGCAAACTGACGCAAGCCAGCCA----- 1912
Oy      1910  CCACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1969
Db      1913  -----CTGCGCGACTCCAAACCCGAGAGTCAATCACTATGAGGAGAAACCCAC 1961
Oy      1970  CGCGTGGCGGAGAACCTGCGGAGGTAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2029
Db      1962  CGCGTGGCGGAGAACCTGCGGAGGTAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2021
Oy      2030  AGGAGCATGCGACGCGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2089
Db      2022  AGGAGCATGCGACGCGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2081
Oy      2090  CATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2147
Db      2082  CATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2141
Oy      2148  AG-CTCGCAGAGAGCATCTCTCTCTCTGTAATCACAGACAGCGGAGCCACAGCGG 2206
Db      2142  AGCTCGCAGAGAGCATCTCTCTCTCTGTAATCACAGACAGCGGAGCCACAGCGG 2201
Oy      2207  CCAAAA-TTTCAGAGCAGAG-TTMAAGATGTGTGTAATGATGTAATGATGTAATGATGTAAT 2264
Db      2202  CCAAAAATTTCAGAGCAGAGTTTCAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2261
Oy      2265  TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2284
Db      2262  TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2281

```

RESULT 15

```

US-10-189-123-33
; Sequence 33, Application US/10189123
; Publication No. US20030082586A1
; GENERAL INFORMATION:
; APPLICANT: KIRST, Susan J.
; APPLICANT: HOLTZMAN, Douglas A.
; APPLICANT: FRASER, Christopher C.
; APPLICANT: SHARP, John D.
; APPLICANT: BARNES, Thomas S.
; TITLE OF INVENTION: ANTIBODIES HAVING DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE REFERENCE: 10147-1103
; CURRENT APPLICATION NUMBER: US/10/189,123
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 09/342,364

```

```

; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 33
; LENGTH: 2811
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-189-123-33

Query Match      52.3% Score 1193.6; DB 14; Length 2811;
Best Local Similarity 74.6% Pred. No. 0; Mismatches 423; Indels 161; Gaps 11;
Matches 1716; Conservative 0;

Oy      118  CGCACCAGCAGAGCAGCTGCGCTGGAAGCTCAGAGCCGCGGCGCTGCGCATGCGCCACAC 177
Db      10  CGCGTGGGAGACAGCTGCGCTGGAAGCTCAGAGCCGCGGCGCTGCGCATGCGCCACAC 69
Oy      178  TGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 237
Db      70  TGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 129
Oy      238  TGAACCTTGTGCGCGTAAAGACAGTGTCTTGAGAGCTGCGCCACGCTAGTACGCGG 297
Db      130  TGGAACTTGTGCGCGTAAAGACAGTGTCTTGAGAGCTGCGCCACGCTAGTACGCGG 189
Oy      238  TTCTCCAGACCGGATCCAGGACTTCTGACACTGACCTGACGAGCGAGCCACTGGGCTT 357
Db      190  TTCTCCAGACCGGATCCAGGACTTCTGACACTGACCTGACGAGCGAGCCACTGGGCTT 249
Oy      358  CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 417
Db      250  CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 309
Oy      418  GGAGCGATCTCTGCGGAGGCGCCGCTGAGAGAAAGACTGATATCCAGAAAGGAG 477
Db      310  GGAGCGATCTCTGCGGAGGCGCCGCTGAGAGAAAGACTGATATCCAGAAAGGAG 369
Oy      478  AACAACAGACCGAGTCTTCAATTCATCGCTCTCTGACAGCCCTACAACTGCTCCAC 537
Db      370  AACAACAGACCGAGTCTTCAATTCATCGCTCTCTGACAGCCCTACAACTGCTCCAC 429
Oy      538  CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 585
Db      430  CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 489
Oy      586  ----- 585
Db      490  CTCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 549
Oy      586  -----AACA 589
Db      550  GGCCCAAGAGCCCTGCGCTTAAAGCATCTCATCACTCTCTGCTGCTTAAACA 609
Oy      590  TGCCTACCTTCACTTGGAGCATGAGAGTTGAAGTGAAGAGGCAAGTCCCTATG 649
Db      610  TGCCTACCTTCACTTGGAGCATGAGAGTTGAAGTGAAGAGGCAAGTCCCTATG 669
Oy      650  ACCAGCTAAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 709
Db      670  ACCAGCTAAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 729
Oy      710  ACAACTTCTGCGGACAGGAACCATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 769
Db      730  ACAACTTCTGCGGACAGGAACCATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 789
Oy      770  AGACAGAGTACTGCGCTTGGCTCAAGCAAGCTCACTTGTAGCTGCTGCTGCTGCTGCTGCT 829
Db      790  AGACAGAGTACTGCGCTTGGCTCAAGCAAGCTCACTTGTAGCTGCTGCTGCTGCTGCTGCT 849
Oy      830  CTGAGAGTGTGGGACGCTTACAGGCGGAGAGCAAGAGTCTACTTCTTCAAGGAGC 889
Db      850  CTGAGAGTGTGGGACGCTTACAGGCGGAGAGCAAGAGTCTACTTCTTCAAGGAGC 909

```

QY	890	GGGAGAGTGAATCCGACGTGCATCCGAGCAGTGGTGGCTGTGGGCCGTGTCTGCA	949
Db	910	GGGAGTGGAGTCCGACTGCTATCCATCCAGCAGAGTGGTGGCTGTGGGCCGTGTCTGCA	969
QY	950	AGGGCGATATGGGGGGGCGCAGCGACCCCTGCAGAGAAATGGACACGTTCCGTGAAGCGC	1009
Db	970	AGGGCGATATATGGGGGGGCGCAGCGACCCCTGCAGAGAAATGGACACGTTCCGTGAAGCGC	1029
QY	1010	GGCTGGCATGTCTGTGCCCCGAACTGGCAGCTCTTCAACCACTGTGAGCGATATGCACA	1069
Db	1030	GGCTGGCATGTCTGTGCCCCGAACTGGCAGCTCTTCAACCACTGTGAGCGATATGCACA	1089
QY	1070	CCCTGCAGAGACCTCCCTGGCAGCAACACACTCTTTGGGGTTTTCACAGCAAGTGGG	1129
Db	1090	CCCTGCAGAGACCTCCCTGGCAGCAACACACTCTTTGGGGTTTTCACAGCAAGTGGG	1149
QY	1130	GTGACATGTACCTGTGGCCATCTGTGTAGTACCAAGTTGGAGAGATCCAGCGGGTGTGG	1189
Db	1150	GTGACATGTACCTGTGGCCATCTGTGTAGTACCAAGTTGGAGAGATCCAGCGGGTGTGG	1209
QY	1190	AGGGCCCCATTAAGAGTACCATGAGAGCCCAAGAAAGTGGACCGCTACACTGACCCGTG	1249
Db	1210	AGGGCCCCATTAAGAGTACCATGAGAGCCCAAGAAAGTGGACCGCTACACTGACCCGTG	1269
QY	1250	TACCCAGCCCTCGGGCCCTGCTGCTGTGTCATTAACAATCGCATCGGGCCGACGGCTACACA	1309
Db	1270	TACCCAGCCCTCGGGCCCTGCTGCTGTGTCATTAACAATCGCATCGGGCCGACGGCTACACA	1328
QY	1310	GCTCCCTGGAGACTACCCGCAACAATCCATCTGTCTCAAGAACACCCCGCTGATGGAGG	1369
Db	1329	TTTTCAGAGAGACAGAGGGGCGCGGCTGTGCTGTGAAGCTTACCTTGTGCTGTGGCAG	1388
QY	1370	AGCAGTGGGGGCTCGGTGGAGCCGCCCTGTCTGTGAAGAAAGGCAACCACTTCACCC	1429
Db	1389	GCCGCTGGGTGACCTTGGAGGGCCGGGGCCCTGTGAAGAAACCTGGGGCTGTGTGGCTGG	1448
QY	1430	ACCTGTCGCGGACCGGGTTACAGAGCTGTATGAGAGCACCTATACATGTGTTTCATTTG	1489
Db	1449	CGGTGTGGCCCTGGGGGCTGTGTGCC---TGTGTGCTGTGTGTGTGTGTGTGTGTGTGTG	1504
QY	1490	GCACAGAGACGCGTGGCTGTCTCAAGGCTGTGACGCTGGGGCCCTGGGTTACCTGATTTG	1549
Db	1505	CGCGGGCGGCTGGCGGGGAAGACTGGAGAAAGGGGCCAAAGCTACTGAGAGGACCT--TGG	1562
QY	1550	AGGAGCTGCAGCTTTTGGACGAGACCCATGAGAAAGCGCTGTGCTATCTCAGAGCAGAA	1609
Db	1563	TGTACCCCTGGAGCTGGCCAAAGAGGCCACACAGTCCGCCCTTCC--GGCCCTGTCTGGA	1620
QY	1610	AGCTGCTTTTGGCGGCTCCCGCTCTACGCTGTGTGACGCTGCCCGTGGCCGACGTGATAA	1669
Db	1621	ACCGAGATGAGAAATTTGGGATCCTGTGCTTACTACTATTCAGATGGCTCCCTTAAAGAT	1680
QY	1670	AGTATGCGTCTGTGAGACGTGTCCCGCCGCGGAGCCCTATTTGGGCTTGGAGCGTCA	1729
Db	1681	AGTATGCGTCTGTGAGACGTGTCCCGCCGCGGAGCCCTATTTGGGCTTGGAGCGCAT	1740
QY	1730	ACACAGCCGCTGTGTGGCCGTGGGTGGCCACTTTTGTGATCTTTACTGATCAGCATGTGA	1789
Db	1741	CCGAGGCGAGCTCTGCTCTTCATCCAGCTCGGCTTCACTGGGGGGTGGCGGAACTCAAA	1800
QY	1790	TGACCTTGGACACTTACAGGACTTTGCAACCTCCGTGGCAGTAAATATACGTACAGGCCAC	1849
Db	1801	TGCCAATGTGTACGTGCGCTTACAACTAGAGGGGAGAGACCGGGAGGGCTCGGGCACCC	1860
QY	1850	TNCCCAAAAATATCAGGTGTGGCGGGGCAAGACGTGTGTCTGCTCCGACACTCTCTCT	1909
Db	1861	CTGTGCTGTGTGGGATGAACTGAGACGGAAACTGCAGCAACGCCACCA-----	1912
QY	1910	CCAACTTGGCCCTGCCGACTCCAAACCCGAGAGTCACTAGTATGAGGGGAACCCCAAC	1969
Db	1913	-----CTGGCGGACTCCAAACCCCGAGAGTATATGATATGAGGGGAACCCCAAC	1961
QY	1970	CGCGTGGCGGAGAAAGCTGGGAGGTGTACTCTTACTTTTGCACAGCCACCAAGTATCTC	2029

Search completed: August 19, 2003, 17:23:14
Job time : 427 secs

Search completed: August 19, 2003, 17:23:14
Job time : 427 secs

THIS PAGE BLANK (USPTO)